

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representation of  
The original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C12N 15/12, 15/55, 9/10, C12Q 1/68,</b> <b>C07K 14/47, A61K 48/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/14234</b> <b>(43) International Publication Date:</b> 16 March 2000 (16.03.00)
<b>(21) International Application Number:</b> PCT/US99/20544 <b>(22) International Filing Date:</b> 7 September 1999 (07.09.99)  <b>(30) Priority Data:</b> 60/099,338 8 September 1998 (08.09.98) US  <b>(71) Applicant:</b> UROCOR, INC. [US/US]; 840 Research Parkway, Oklahoma City, OK 73104-3699 (US).  <b>(72) Inventors:</b> AN, Gang; 2609 N.W. 160th Street, Edmond, OK 73013 (US). VELTRI, Robert; 7301 N.E. 106th Street, Oklahoma City, OK 73151 (US).  <b>(74) Agent:</b> NAKASHIMA, Richard, A.; Arnold White & Durkee, P.O. Box 4433, Houston, TX 77210 (US).	<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>	
<b>(54) Title:</b> PROSTATE SPECIFIC PROMOTER AND REGULATION OF GENE EXPRESSION  <b>(57) Abstract</b>  Disclosed are compositions and methods of use of the promoter for prostate specific transglutaminase (SEQ ID NO:15). The invention relates particularly to isolated nucleic acids and vectors comprising the sequence of this promoter. The invention also relates to methods of therapeutic treatment for prostate cancer or benign prostatic hyperplasia (BPH) utilizing this promoter. Described are means for the isolation and identification of transcriptional factors and other DNA-binding proteins that regulate promoter transcriptional activity, identification of regulatory elements within the promoter and construction of deletion mutants containing specific subsets of these regulatory elements, identification of small molecule ligands that bind to and inhibit or activate the identified transcriptional factors and other DNA-binding proteins, construction of vectors containing the prostate specific transglutaminase promoter operatively linked to genes of use in the treatment of prostate cancer or BPH, and methods for treatment of prostate cancer or BPH by administration of such vectors to patients with prostate cancer or BPH. Further described are methods for treatment of prostate cancer or BPH by administration of small molecule ligands that bind to and inhibit or activate transcriptional factors or other DNA-binding proteins that regulate the activity of this promoter.		

BH

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## **DESCRIPTION**

### **PROSTATE SPECIFIC PROMOTER AND REGULATION OF GENE EXPRESSION**

#### **1.0 BACKGROUND OF THE INVENTION**

This application claims priority under 35 U.S.C. § 119(e) to provisional application  
5 Serial No. 60/099,338, filed on September 8, 1998. The entire text of the above-referenced  
application is specifically incorporated herein by reference without disclaimer.

#### **1.1 Field of the Invention**

The present invention relates generally to the fields of regulation of gene expression in  
10 tissues of prostate origin using a novel prostate specific promoter and methods relating thereto.  
More particularly, the present invention concerns compositions comprising prostate specific  
promoter sequences and methods useful in regulating gene expression in tissues of prostate origin,  
including metastatic and non-metastatic prostate cancers. Also disclosed are various therapeutic  
methods using the compositions of the invention.

15

#### **1.2 Description of the Related Art**

Carcinoma of the prostate (PCA) is the most frequently diagnosed cancer among men in  
the United States (Veltri *et al.*, 1996). Prostate cancer was diagnosed in approximately 189,500  
men in 1998 and about 40,000 men succumbed to the malignancy (Landis *et al.*, 1998).  
20 Unfortunately, the relatively few prostate carcinomas that are progressive in nature are likely to  
have already metastasized by the time of clinical detection. Survival rates for individuals with  
metastatic prostate cancer are quite low.

The standard therapy for individuals with advanced metastatic prostate cancer is total  
androgen ablation by means of castration, antiandrogens or luteinizing hormone-releasing  
25 hormone analogs (Konety and Getzenberg, 1997). However, in most patients, the metastatic  
tumors progress to hormone refractory disease within two to three years, resulting in patient  
death (Konety and Getzenberg, 1997). Currently available methods of treatment, such as  
chemotherapy or radiation therapy, are of limited efficacy for such individuals (Hrouda and  
Dalgleish, 1996; Goethuys *et al.*, 1997).

30 Thus, a need exists for the development of new methods of treatment targeted to prostate  
cancer cells, particularly those that have progressed to metastatic, hormone refractory disease.  
One approach to this problem consists of attaching cytotoxic or cytostatic genes to prostate  
specific promoters, followed by gene therapy. Although a number of studies have demonstrated

the potential feasibility of this approach in cell culture and animal model systems (Malkowicz and Johnson, 1998), the development of effective methods of gene therapy for treatment of advanced prostate cancers in humans is a long-standing need in the field of cancer therapy (Segawa *et al.*, 1998).

5       The majority of studies in this field to date have utilized the promoter and enhancer for prostate specific antigen (PSA). Martiniello-Wilks *et al.* (1998) incorporated genes encoding either HSVTK (herpes simplex virus-thymidine kinase) or PNP (purine nucleoside phosphorylase) into attenuated human adenovirus (Ad5) under the control of a PSA promoter. Injection of the recombinant Ad5 into PC3 (human prostate cancer) tumors in nude mice was followed by  
10   treatment with the prodrugs GCV (gancyclovir) for HSVTK and 6MPDR (6-methylpurine deoxyriboside) for PSP. Growth of PC3 tumors was reportedly inhibited 80% by HSVTK/GCV treatment and 75% by PNP/6MPDR treatment, associated with an increase in animal survival (Martiniello-Wilks *et al.*, 1998). The treatment of solid tumors, including prostate cancer, by administration of a recombinant adenoviral vector containing HSVTK, followed by gancyclovir  
15   administration, was reported by Woo *et al.* (U.S. Patent No. 5,631,236, May 20, 1997). Gotoh *et al.* (1998) also inserted a TK gene under the control of a PSA promoter into adenovirus. Upon infection with the recombinant adenovirus followed by treatment with acyclovir, an androgen-independent subline of LNCaP cells reportedly showed increased cell death and an inhibition of tumor growth in castrated animals (Gotoh *et al.*, 1998).

20       Segawa *et al.* (1998) used a PSA promoter to express a GAL4-VP16 fusion protein. Gal4-responsive elements were incorporated into a separate gene construct encoding polyglutamine, amplifying the activity of the PSA promoter. The resultant polyglutamine expression induced apoptosis in prostate cancer cells (Segawa *et al.*, 1998). The PSA gene enhancer/promoter, attached to a luciferase reporter gene, has been used for screening  
25   therapeutic agents for treatment of prostate cancer (Lamparski *et al.*, U.S. Patent No. 5,783,435, July 21, 1998). Other examples of prostate tissue-specific promoters have been reported, including the promoters for probasin (Yan *et al.*, 1997; Matusik, U.S. Patent No. 5,783,681, July 21, 1998) and prostatic acid phosphatase (Zelivianski *et al.*, 1998).

Although PSA is a prostate tissue-specific promoter, it apparently is not differentially  
30   regulated in prostate cancer cells compared with normal prostate cells or benign prostatic hyperplasia. (Croce *et al.*, WO 94/10343, 1994; Ghossein *et al.*, 1995; Gomella *et al.*, 1997; Olssen *et al.*, 1997). In fact, "Within the prostate PSA expression is independent of the degree of dysplasia since prostatic cells from normal, benign hyperplastic, malignant and metastatic

tissue express and secrete PSA." (Gomella *et al.*, 1997). This characteristic of PSA as a prostate tissue-specific promoter, rather than a prostate cancer-specific promoter, indicates that PSA-promoter linked gene therapies will target normal prostate tissues as well as prostate cancers. For the maximum therapeutic benefit to be derived from gene therapy treatments, it is desirable that promoters be identified for genes that are differentially expressed in prostate cancer cells compared with normal prostate tissues. The instant invention addresses this need through the identification, characterization and methods of use of a novel, prostate specific promoter that is differentially expressed in prostate cancer compared with normal prostate tissues, associated with the gene encoding prostate specific transglutaminase.

## 2.0 SUMMARY OF THE INVENTION

The present invention addresses deficiencies in the prior art by providing a novel method of gene therapy for prostate cancer in humans, using the promoter for prostate specific transglutaminase operably linked to specific genes. Such genes may include tumor suppressor genes (P53), suicide genes (cytosine deaminase, thymidine kinase, PNP, nitroreductase, diphtheria toxin), anti-oncogenes (anti-sense c-myc), cell adhesion genes (E-cadherin), genes encoding catalytic antisense RNA (anti-bcl-2 ribozyme), cytokine genes and apoptosis genes (polyglutamine) (Malkowicz and Johnson, 1998).

The prostate specific transglutaminase promoter can also be operably linked to genes capable of activating an immune response against the target prostate cancer cells. Such prostate-specific expression vectors may function as genetic vaccines, according to Weiner and Kennedy (1999). Although any highly antigenic, surface expressed protein or peptide may be of use in the practice of this invention, the skilled artisan will realize that those proteins or peptides that do not provoke an autoimmune response will be preferred. Representative genetic vaccines that could be expressed from a prostate specific transglutaminase promoter include hepatitis B surface antigen, herpes simplex glycoprotein, HIV envelope proteins, hemagglutinin, malaria circumsporozoite protein, carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveolin, POV1, HER2/neu or p27KIP1 (Weiner and Kennedy, 1999; Tjoa *et al.*, 1999; Murphy *et al.*, 1999; Reiter *et al.*, 1998; Yang *et al.*, 1998; Cole *et al.*, 1998; Cordon-Cardo *et al.*, 1998). Co-expression of cytokine genes, such as IL12, may facilitate the induction of an immune response (Weiner and Kennedy, 1999). A partial listing of representative cytokine genes that could be used in the practice of the present invention is included in Table 3 below.

The prostate specific transglutaminase promoter could also be used to drive the expression of humanized antibodies targeted against cancer-associated proteins. An example of a humanized antibody that could be expressed from the prostate specific transglutaminase promoter is the Herceptin® antibody, targeted against the HER2/neu receptor (U.S. Patent No. 5,821,337, incorporated herein by reference). Production and secretion of such proteins in close proximity to prostate cancer cells may be of use in promoting immune responses to the tumor cells. The skilled artisan will realize that such humanized antibodies are not limited to those specific for the HER2/neu receptor, but may also include antibodies against other antigenic proteins or peptides present on the cell surface of prostate cancer cells.

10 An unexpected aspect of the present invention is that the prostate specific transglutaminase promoter is differentially transcribed in prostate cancer cells compared with normal human prostate tissue, allowing for the possibility of differential expression of cytotoxic and other genes in normal prostate compared to prostate cancer. The skilled artisan will realize that the genes listed above are provided only as examples of those that may be useful for prostate cancer gene therapy and that many other known genes would be included within the scope of the instant invention. It is contemplated that any potentially cytocidal or cytostatic gene might be operably linked to the prostate specific transglutaminase promoter or to other promoters identified using the methods disclosed herein. Such promoters operably linked to such genes may be incorporated into vectors and used for prostate cancer therapy within the scope of the present invention.

20 Throughout this specification, the terms promoter and promoter region are used in their broadest sense. Promoters and promoter regions may be generally identified in terms of their functional effect upon the transcription of associated genes. A promoter or promoter region may be broadly considered to comprise a sequence of DNA, either contiguous or non-contiguous, that is effective to provide for the transcription of a gene that is operably linked to the promoter or promoter region. The skilled artisan will realize that a promoter or promoter region may comprise one or more regulatory elements, such as TATA boxes, CAAT boxes, Hogness boxes, GC boxes, and known binding sites for transcriptional factors or DNA polymerases. Such elements may be arranged within a promoter in various combinations and various spatial relationships, including different orientations and different distances from the transcription start site. So long as any such combination is effective to provide for the transcription of an operably linked gene, it may be considered to fall within the definition of a promoter or promoter region. A promoter or promoter region is operably linked to a gene when

it is associated with the gene in an orientation and topological relationship such that it promotes the transcription of the gene.

Specific embodiments of the present invention include isolated nucleic acids, either single- or double-stranded, comprising a prostate specific transglutaminase promoter, compositions comprising isolated nucleic acids having the sequence of SEQ ID NO:15, vectors comprising a prostate specific transglutaminase promoter operably linked to specific genes and vectors comprising deletion mutants of a prostate specific transglutaminase promoter operably linked to specific genes. In one broad aspect, the present invention comprises an isolated nucleic acid of between 20 and 1399 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO:15. In another broad aspect, the present invention comprises an expression vector comprising a prostate specific transglutaminase promoter operably linked to a selected gene, wherein the promoter comprises between 20 and 1399 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO:15.

Vectors that may be used include, but are not limited to, plasmid vectors, naked DNA, viral vectors, including retroviral and DNA vectors, such as adenovirus, adeno-associated virus, vaccinia virus, Sindbis virus, cytomegalovirus, herpes simplex virus, defective hepatitis B viruses, and any other vector or vector system described herein or known in the art. Vectors may be transfected into host cells by means including, but not limited to, viral infection, calcium phosphate precipitation, DEAE-dextran, electroporation, direct microinjection, DNA-loaded liposomes and lipofectamine-DNA complexes, cell sonication, gene bombardment using high velocity microprojectiles, polycations, and receptor-mediated transfection, or any other means described herein or known in the art. Methods of treatment also may include administering modulators of prostate specific transglutaminase transcription, translation, stability or activity.

Further embodiments of the present invention include methods of identifying novel prostate specific promoters, methods of identifying protein binding factors for the prostate specific transglutaminase promoter and methods of identifying regulatory sequences within the prostate specific transglutaminase promoter.

Novel prostate specific promoters may be identified by screening genomic human libraries with probes or primers designed to bind under high stringency conditions with polynucleotides of sequences identical to or complementary to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:14 or SEQ ID NO:15. High stringency conditions are understood to mean conditions under which the probe specifically hybridizes to a target sequence in an



amount that is detectably stronger than non-specific hybridization. High stringency conditions, then, would be conditions which would distinguish a polynucleotide with an exact complementary sequence, or one containing only a few scattered mismatches from a random sequence that happened to have a few small regions (3-10 bases, for example) that matched the probe. Such small regions of complementarity are more easily melted than a full length complement of 14-17 or more bases and high stringency hybridization makes them easily distinguishable.

Relatively high stringency conditions would include, for example, low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl or the equivalent, at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for detecting expression of specific prostate disease markers. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide.

Methods of identifying protein binding factors for the prostate specific transglutaminase promoter may be accomplished by techniques well known in the art, such as affinity chromatography, using double-stranded DNA comprising part or all of the sequence of SEQ ID NO:15 as an affinity ligand. Similar methods may be applied to identify protein binding factors for the semenogelin II promoter, using SEQ ID NO:14 as an affinity ligand. Identification of those regulatory proteins that bind to the prostate specific transglutaminase promoter, but not to the semenogelin II promoter, may serve to identify those factors that are responsible for differential expression of the prostate specific transglutaminase gene in prostate cancer compared to normal prostate tissue.

The identification of regulatory sequences within a promoter may be accomplished by means well known in the art, including construction of promoter deletion mutants operably linked to reporter genes and assay of reporter gene activity. Alternatively, regulatory sequences may be identified by homology with binding sites for known protein regulatory factors or with known enhancer sequences. It is expected that known regulatory elements from other promoters may be incorporated into the prostate specific transglutaminase promoter sequence to change the activity or specificity of the promoter. It is further expected that identified regulatory sequences within the prostate specific transglutaminase promoter may be deleted or modified to change the activity or specificity of the promoter. Such modified promoters may

then be operably linked to specific genes and used for prostate cancer therapy within the scope of the present invention.

Another embodiment of the present invention includes methods of treating individuals with prostate cancer by providing activators or inhibitors of regulatory proteins for prostate specific promoters. An additional embodiment includes methods of treating individuals with prostate cancer by providing an effective amount of a eukaryotic expression vector, comprising a prostate specific transglutaminase promoter operably linked to a selected gene, such as prostate specific transglutaminase. A further embodiment includes providing such a vector wherein the gene encodes a protein that activates a pro-drug, such as gancyclovir or 5-fluorouracil, followed by administration of an effective amount of the pro-drug to the individual with prostate cancer. Yet another embodiment includes a compound for treating individuals with prostate cancer, produced by a method of identifying inhibitors or activators of regulatory proteins that specifically bind to the promoter of prostate specific transglutaminase.

The present invention also provides methods of treating prostate disease, comprising administering to such a patient with prostate disease a therapeutically effective amount of a pharmaceutically acceptable solution containing a composition comprising a prostate specific transglutaminase promoter operably linked to a specific gene. These treatments may comprise administering a composition containing recombinant vectors that express prostate specific transglutaminase proteins or peptides. Such vectors may be administered to a subject *in vivo*, *i.e.* through intravenous administration, or *ex vivo* by transfection into isolated cells that are cultured and then infused into the subject. Such cells are preferably homologous cells, *i.e.*, derived from tissue or serum of the patient, or they may include heterologous cells.

In another broad aspect, the present invention concerns methods for inducing an immune response against prostate cancer cells. Such a response may be induced, for example, by administering to an individual with prostate cancer a therapeutically effective amount of a vector comprising a prostate specific promoter operably linked to an antigenic, surface expressed protein, such as hepatitis B surface antigen, herpes simplex glycoprotein, HIV envelope proteins, hemagglutinin, malaria circumsporozoite protein, carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveollin, POV1, HER2/neu or p27KIP1. Alternatively, an immune response may be induced by administering a vector comprising a prostate specific promoter operably linked to a humanized antibody, such as the Herceptin® antibody.

An aspect of the present invention is a cell-based assay for identifying compounds which affect prostate specific transglutaminase production. Specifically, the assay comprises culturing a cell containing an expression vector comprising a DNA sequence encoding a prostate specific transglutaminase promoter operatively linked to a reporter gene under  
5 conditions which permit expression and quantitative assay of the reporter gene. The cultured cell is incubated with compounds suspected of possessing regulatory activity for production of prostate specific transglutaminase. These regulatory compounds are identified by their ability to modulate the expression of the reporter gene and thereby affect the production of the assayable product of the reporter gene. In certain aspects of the invention the terms  
10 "modulation", "modulate", "affect", "regulate", and "alter" may mean an increase or decrease in the expression of a gene or a gene product's activity.

A particularly useful cell population to use in screening for prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II expression is human tumor cells. Most notably, the present invention is useful in screening compounds which affect prostate specific  
15 transglutaminase and/or cytokeratin 15 production in prostate cancer cells. The present invention is also useful in screening compounds which affect semenogelin II production in lymphocyte cancer cells. A useful prostate cancer cell population in which to perform screening is LNCaP prostate cancer cell line. Other preferred cell lines include DU145, PC-3, C4-2, C4-2Ln and C4-2B (Chung *et al.*, 1994.)

### 3.0 BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed  
5 description of specific embodiments presented herein.

FIG. 1. Nucleic acid sequence of the promoter for prostate specific transglutaminase (SEQ ID NO:15). The sequence of the prostate specific transglutaminase promoter, identified as described in Example 2, is presented according to standard nomenclature, with the first  
10 transcribed base indicated as position +1 and the 5' untranscribed sequence, including the promoter sequence, indicated from -1400 to -1 (corresponding to positions 1 to 1400 of SEQ ID NO:15). The 5' transcribed region is indicated from +1 to +53 (corresponding to positions 1401 to 1453 of SEQ ID NO:15).

15 FIG. 2. CAT assay for the activity of the 1.4 kb prostate specific transglutaminase promoter. The lower part of FIG. 2 shows the 1.4 kb insert into the pCAT3 vector. The upper part of FIG. 2 shows a TLC assay of CAT activity, using a radiolabeled acetylated CAT product. Ref. = product reference; Neg. = negative control; Pos. = positive control; PC3 = prostate cancer cell line; T24 = bladder cancer cell line.

20 FIG. 3. CAT assay for the activity of the 0.8 kb (5') prostate specific transglutaminase promoter. The lower part of FIG. 3 shows the 0.8 kb 5' promoter insert into the pCAT3 vector. The upper part of FIG. 3 shows a TLC assay of CAT activity, using a radiolabeled acetylated CAT product. Ref. = product reference; Neg. = negative control; Pos. = positive control; PC3 =  
25 prostate cancer cell line; T24 = bladder cancer cell line.

### 4.0 DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention concerns compositions and methods of use of the promoter region for prostate specific transglutaminase (SEQ ID NO:15). More specifically, the invention concerns  
30 isolated nucleic acids and vectors comprising the sequence of this promoter region. Methods of therapeutic treatment for prostate cancer or benign prostatic hyperplasia (BPH) utilizing this promoter region also are within the scope of the invention. Such methods comprise the isolation and identification of transcriptional factors and other DNA-binding proteins that

regulate promoter region transcriptional activity, identification of regulatory elements within the promoter region and construction of deletion mutants containing specific subsets of these regulatory elements, identification of small molecule ligands that can bind to and inhibit or activate the identified transcriptional factors or other DNA-binding proteins, construction of  
5 expression vectors containing the prostate specific transglutaminase promoter region operatively linked to genes of use in the treatment of prostate cancer or BPH, and administration of such vectors to patients with prostate cancer or BPH.

Those skilled in the art will realize that the nucleic acid sequences disclosed will find utility in a variety of applications in the treatment of prostate disease. Examples of such  
10 applications within the scope of the present invention comprise therapeutic treatments of prostate disease using sense or antisense expression vectors, expression activators or inhibitors specific for the identified promoter sequence, genetic vaccines and humanized antibodies.

#### 4.1 Nucleic Acids

As described herein, an aspect of the present disclosure is three markers of prostate  
15 disease, identified by Southern differential hybridization. Northern analysis, and quantitative RT-PCR™. These include the nucleic acid products of prostate specific transglutaminase (GenBank accession #s L34840 and I20492), cytokeratin 15 (GenBank accession # X07696), and semenogelin II (GenBank accession #s M81652 and M81651). The under-expression of these  
20 gene products in metastatic prostate cancer was first reported in co-pending patent application U.S.S.N. 09/010,398, the entire text of which is incorporated herein by reference.

In one embodiment, the nucleic acid sequences disclosed herein will find utility as hybridization probes or amplification primers. These nucleic acids may be used, for example, in diagnostic evaluation of tissue or serum samples. In certain embodiments, these probes and  
25 primers consist of oligonucleotides. Such oligonucleotides are of sufficient length to provide specific hybridization to a RNA or DNA target derived from a tissue or serum sample, typically of 10-20 nucleotides, but in some cases longer. Longer sequences, *e.g.*, 30, 40, 50, 100, 500 nucleotides and even up to full length, as disclosed in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:14 or SEQ ID NO:15 are preferred for certain embodiments.

30 Nucleic acid molecules having contiguous stretches of about 10, 15, 17, 20, 30, 40, 50, 60, 75 or 100 or 500 nucleotides homologous to a sequence selected from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:14 or SEQ ID NO:15 are contemplated. Molecules that bind to these sequences under high stringency conditions also are contemplated. These probes will be

useful in a variety of hybridization embodiments, such as Southern blotting, Northern blotting and *in situ* hybridization. In some cases, it is contemplated that probes may be used that hybridize to multiple target sequences without compromising their utility.

Various probes and primers can be designed around the disclosed nucleotide sequences.

- 5 Primers may be of any length but, typically, are 10-20 bases in length. By assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all primers can be proposed:

$$n \text{ to } n + y$$

- 10 where  $n$  is an integer from 1 to the last number of the sequence and  $y$  is the length of the primer minus one (9 to 19), where  $n + y$  does not exceed the last number of the sequence. Thus, for a 10-mer, the probes correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the probes correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the probes correspond to bases 1 to 20, 2 to 21, 3 to 22 ... and so on. For the sequences disclosed herein,  $n = 3064$  (SEQ ID NO:1);  $n = 1709$  (SEQ ID NO:2);  $n = 1981$  (SEQ ID NO:3);  $n = 8224$  (SEQ ID NO:14); and  $n =$   
15 1453 (SEQ ID NO:15).

- The use of a hybridization probe of between 14 and 100 nucleotides or longer in length allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 20 bases in length are generally preferred, in  
20 order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of particular hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having stretches of 20 to 30 nucleotides, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means or by introducing selected sequences into recombinant vectors for recombinant production.

- 25 Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of genes or RNAs or to provide primers for amplification of RNA from tissue or serum. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence.

- 30 For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch

between the probe and the template or target strand, and would be particularly suitable for isolating specific genes or detecting specific mRNA transcripts. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide.

For certain applications, for example, substitution of amino acids by site-directed mutagenesis, it is appreciated that lower stringency conditions are required. Under these conditions, hybridization may occur even though the sequences of probe and target strand are not perfectly complementary, but are mismatched at one or more positions. Conditions may be rendered less stringent by increasing salt concentration and decreasing temperature. For example, a medium stringency condition could be provided by about 0.1 to 0.25 M NaCl at temperatures of about 37°C to about 55°C, while a low stringency condition could be provided by about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

The following codon chart may be used, in a site-directed mutagenic scheme, to produce nucleic acids encoding the same or slightly different amino acid sequences of a given nucleic acid:

**TABLE 1. Codon Usage**

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			

Table 1 – Continued

Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In other embodiments, hybridization may be achieved under conditions of, for example, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, at temperatures between approximately 20°C to about 37°C. Other hybridization conditions utilized could include approximately 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, at temperatures ranging from approximately 40°C to about 72°C.

In certain embodiments, it will be advantageous to employ nucleic acid sequences of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of being detected. In preferred embodiments, one may desire to employ a fluorescent label or an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmentally undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a detection means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

In general, it is envisioned that the hybridization probes described herein will be useful both as reagents in solution hybridization, as in PCR™, for detection of expression of corresponding genes, as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to hybridization with selected probes under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required (depending, for example, on the G+C content, type of target nucleic acid, source of nucleic acid, size of hybridization probe, etc.).



Following washing of the hybridized surface to remove non-specifically bound probe molecules, hybridization is detected, or even quantified, by means of the label.

It is understood that this disclosure is not limited to the particular probes disclosed herein and particularly is intended to encompass at least isolated nucleic acids that are hybridizable to nucleic acids comprising the disclosed sequences or that are functional sequence analogs of these  
5 nucleic acids. For example, a nucleic acid of partial sequence may be used to quantify the expression of a structurally-related gene or the full length genomic or cDNA clone from which it is derived.

For applications in which the nucleic acid segments of the present invention are  
10 incorporated into vectors, such as plasmids, cosmids or viruses, these segments may be combined with other DNA sequences, such as promoter regions, polyadenylation signals, restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the  
15 intended recombinant DNA protocol.

DNA segments encoding a specific gene may be introduced into recombinant host cells and employed for expressing a specific structural or regulatory protein. Alternatively, through the application of genetic engineering techniques, subportions or derivatives of selected genes may be employed. Upstream regions containing regulatory regions such as promoter region regions may  
20 be isolated and employed for expression of the selected gene.

#### 4.2 Encoded Proteins

The metastatic cancer marker genes described herein can be inserted and expressed in any number of different recombinant DNA expression systems to generate large amounts of the  
25 polypeptide product, which can then be purified and used to vaccinate animals to generate antisera for use in the practice of the present invention.

The gene or gene fragment encoding a polypeptide may be inserted into an expression vector by standard subcloning techniques. An *E. coli* expression vector may be used which produces the recombinant polypeptide as a fusion protein, allowing rapid affinity purification of  
30 the protein. Examples of such fusion protein expression systems are the glutathione S-transferase system (Pharmacia, Piscataway, NJ), the maltose binding protein system (NEB, Beverly, MA), the FLAG system (IBI, New Haven, CT), and the 6xHis system (Qiagen, Chatsworth, CA).

#### 4.2.1 Expression of Proteins from Cloned cDNAs

The cDNA species specified in SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3 may be expressed as encoded peptides or proteins. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. It is believed that virtually any expression system may be employed in the expression of the isolated cDNA species or the nucleic acid sequences for the disclosed prostate disease marker genes.

Both cDNA and genomic sequences are suitable for eukaryotic expression, as the host cell will generally process the genomic transcripts to yield functional mRNA for translation into protein. In addition, it is possible to use partial sequences for generation of antibodies against discrete portions of a gene product, even when the entire sequence of that gene product remains unknown. Computer programs are available to aid in the selection of regions which have potential immunologic significance. For example, software capable of carrying out this analysis is readily available commercially, for example MacVector (IBI, New Haven, CT). The software typically uses standard algorithms such as the Kyte/Doolittle or Hopp/Woods methods for locating hydrophilic sequences which are characteristically found on the surface of proteins and are, therefore, likely to act as antigenic determinants.

As used herein, the terms "engineered" and "recombinant" cells are intended to refer to a cell into which an exogenous DNA segment or gene, such as a cDNA or gene has been introduced through the hand of man. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced exogenous DNA segment or gene. Recombinant cells include those having an introduced cDNA or genomic gene, and also include genes positioned adjacent to a heterologous promoter region not naturally associated with the particular introduced gene. The heterologous gene may be inserted into the host genome or maintained on an episome.

To express a recombinant encoded protein or peptide, whether mutant or wild-type, in accordance with the present invention one would prepare an expression vector that comprises one of the claimed isolated nucleic acids under the control of, or operatively linked to, one or more promoter regions. To bring a coding sequence "under the control of" a promoter region, one positions the 5' end of the transcription initiation site of the transcriptional reading frame generally between about 1 and about 50 nucleotides "downstream" (*i.e.*, 3') of the chosen promoter region.

The "upstream" promoter region stimulates transcription of the DNA and promotes expression of the encoded recombinant protein. This is the meaning of "recombinant expression" in this context.

Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order to achieve protein or peptide expression in a variety of host-expression systems. Cell types available for expression include, but are not limited to, bacteria, such as *E. coli* and *B. subtilis* transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors.

In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is often transformed using pBR322, a plasmid derived from an *E. coli* species. pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoter regions which may be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism may be used as transforming vectors in connection with these hosts. For example, the phage lambda GEM<sup>TM</sup>-11 may be utilized in making a recombinant phage vector which may be used to transform host cells, such as *E. coli* LE392.

Further useful vectors include pIN vectors (Inouye *et al.*, 1985); and pGEX vectors, for use in generating glutathione S-transferase (GST) soluble fusion proteins for later purification and separation or cleavage. Other suitable fusion proteins are those with  $\beta$ -galactosidase, ubiquitin, or the like.

Promoter regions that are most commonly used in recombinant DNA construction include the  $\beta$ -lactamase (penicillinase), lactose and tryptophan (trp) promoter region systems. While these are the most commonly used, other microbial promoter regions have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling those of skill in the art to ligate them functionally with plasmid vectors. For specific purposes of the present invention, it may be desirable to use the promoter regions of a prostate specific gene, such as the promoters for prostate specific transglutaminase (SEQ ID NO:15) or semenogelin II (SEQ ID NO:14).

In addition to microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or

invertebrate culture. In addition to mammalian cells, these include insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus); and plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid)  
5 containing one or more coding sequences.

Examples of useful mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, W138, BHK, COS-7, 293, HepG2, 3T3, RIN and MDCK cell lines. In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications  
10 (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the encoded protein.

Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cells lines or host systems may be chosen to ensure the correct modification and processing of the foreign protein expressed. Expression  
15 vectors for use in mammalian cells ordinarily include an origin of replication (as necessary), a promoter region located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences. The origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (*e.g.*, Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the  
20 vector is integrated into the host cell chromosome, the latter is often sufficient.

A number of viral based expression systems may be utilized, for example, commonly used promoter regions are derived from polyoma, Adenovirus 2, and most frequently Simian Virus 40 (SV40). The early and late promoter regions of SV40 virus are particularly useful because both  
25 are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the Hind III site toward the Bgl I site located in the viral origin of replication.

In cases where an adenovirus is used as an expression vector, the coding sequences may be  
30 ligated to an adenovirus transcription/ translation control complex, *e.g.*, the late promoter region and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*,

region E1 or E3) will result in a recombinant virus that is viable and capable of expressing proteins in infected hosts.

Specific initiation signals may also be required for efficient translation of the claimed isolated nucleic acid coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may need to be provided. One of ordinary skill in the art would readily be capable of determining this and providing the necessary signals. It is well known that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons may be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements or transcription terminators (Bittner *et al.*, 1987).

In eukaryotic expression, one will also typically desire to incorporate into the transcriptional unit an appropriate polyadenylation site (*e.g.*, 5'-AATAAA-3') if one was not contained within the original cloned segment. Typically, the poly A addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express constructs encoding proteins may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells may be transformed with vectors controlled by appropriate expression control elements (*e.g.*, promoter region, enhancer, sequences, transcription terminators, polyadenylation sites, *etc.*), and a selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn may be cloned and expanded into cell lines.

A number of selection systems may be used, including but not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska *et al.*, 1962) and adenine phosphoribosyltransferase genes (Lowy *et al.*, 1980), in tk-, hgp<sup>r</sup>t- or ap<sup>r</sup>t- cells, respectively. Also, antimetabolite resistance may be used as the basis of selection for dhfr, that confers resistance to methotrexate (Wigler *et al.*, 1980; O'Hare *et al.*, 1981); gpt, that confers resistance to mycophenolic acid (Mulligan *et al.*,

1981); neo, that confers resistance to the aminoglycoside G-418 (Colberre-Garapin *et al.*, 1981); and hygro, that confers resistance to hygromycin (Santerre *et al.*, 1984).

It is contemplated that the isolated nucleic acids of the invention may be "overexpressed", *i.e.*, expressed in increased levels relative to its natural expression in human prostate cells or peripheral blood cells, or even relative to the expression of other proteins in the recombinant host cell. Such overexpression may be assessed by a variety of methods, including radio-labeling and/or protein purification. However, simple and direct methods are preferred, for example, those involving SDS/PAGE and protein staining or Western blotting, followed by quantitative analyses, such as densitometric scanning of the resultant gel or blot. A specific increase in the level of the recombinant protein or peptide in comparison to the level in natural human prostate cells is indicative of overexpression, as is a relative abundance of the specific protein in relation to the other proteins produced by the host cell and, *e.g.*, visible on a gel.

#### 4.2.2 Purification of Proteins

Further aspects of the present invention concern the purification, and in particular embodiments, the substantial purification, of proteins or peptides. The term "purified protein or peptide" as used herein, is intended to refer to a composition, isolatable from other components, wherein the protein or peptide is purified to any degree relative to its naturally-obtainable state, *i.e.*, in this case, relative to its purity within a prostate cell extract. A purified protein or peptide therefore also refers to a protein or peptide, free from the environment in which it may naturally occur.

Generally, "purified" will refer to a protein or peptide composition which has been subjected to fractionation to remove various other components, and which composition substantially retains its expressed biological activity. Where the term "substantially purified" is used, this will refer to a composition in which the protein or peptide forms the major component of the composition, such as constituting about 50% or more of the proteins in the composition.

Various methods for quantifying the degree of purification of the protein or peptide will be known to those of skill in the art in light of the present disclosure. These include, for example, determining the specific activity of an active fraction, or assessing the number of polypeptides within a fraction by SDS/PAGE analysis. A preferred method for assessing the purity of a fraction is to calculate the specific activity of the fraction, to compare it to the specific activity of the initial extract, and to thus calculate the degree of purity, herein assessed by a "-fold purification number". The actual units used to represent the amount of activity will, of course, be dependent upon the

particular assay technique chosen to follow the purification and whether or not the expressed protein or peptide exhibits a detectable activity.

Various techniques suitable for use in protein purification will be well known to those of skill in the art. These include, for example, precipitation with ammonium sulphate, PEG, antibodies and the like or by heat denaturation, followed by centrifugation; chromatography steps such as ion exchange, gel filtration, reverse phase, hydroxylapatite and affinity chromatography; isoelectric focusing; gel electrophoresis; and combinations of such and other techniques. As is generally known in the art, it is believed that the order of conducting the various purification steps may be changed, or that certain steps may be omitted, and still result in a suitable method for the preparation of a substantially purified protein or peptide.

There is no general requirement that the protein or peptide always be provided in the most purified state. Indeed, it is contemplated that less substantially purified products will have utility in certain embodiments. Partial purification may be accomplished by using fewer purification steps in combination, or by utilizing different forms of the same general purification scheme. For example, it is appreciated that a cation-exchange column chromatography performed utilizing an HPLC apparatus will generally result in a greater -fold purification than the same technique utilizing a low pressure chromatography system. Methods exhibiting a lower degree of relative purification may have advantages in total recovery of protein product, or in maintaining the activity of an expressed protein.

The migration of a polypeptide may vary, sometimes significantly, with different conditions of SDS/PAGE (Capaldi *et al.*, 1977). It will therefore be appreciated that under differing electrophoresis conditions, the apparent molecular weights of purified or partially purified expression products may vary.

#### 4.3 Preparation of Antibodies Specific for Encoded Proteins

For some embodiments, it will be desirable to produce antibodies that bind with high specificity to the polypeptide product(s) of an isolated nucleic acid selected from SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3 or the disclosed prostate disease marker genes: prostate specific transglutaminase, cytokeratin 15, and semenogelin II. Means for preparing and characterizing antibodies are well known in the art (See, *e.g.*, Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988; incorporated herein by reference).

Monoclonal antibodies (MAbs) may be readily prepared through use of well-known techniques, such as those exemplified in U.S. Patent 4,196,265, incorporated herein by reference.

Typically, this technique involves immunizing a suitable animal with a selected immunogen composition, *e.g.*, a purified or partially purified expressed protein, polypeptide or peptide. The immunizing composition is administered in a manner effective to stimulate antibody producing cells.

5           It will be appreciated by those of skill in the art that monoclonal or polyclonal antibodies specific for proteins that are preferentially expressed in metastatic or nonmetastatic human prostate cancer or prostate disease will have utility in several types of applications. These may include the production of humanized antibodies for use in the induction of an immune response. Antibodies may also be packaged into diagnostic kits for use in detecting or diagnosing human prostate  
10       disease. An alternative use would be to link such antibodies to therapeutic agents, such as chemotherapeutic agents, followed by administration to individuals with prostate disease, thereby selectively targeting the prostate disease cells for destruction. The skilled practitioner will realize that such uses are within the scope of the present invention.

#### 15       4.4     Design of Humanized Antibodies

          In certain applications, genes encoding humanized antibody polypeptides may be designed for use in expression vectors, wherein the genes are operably linked to a prostate specific promoter. A general scheme for design and production of humanized antibodies is presented in U.S. Patent No. 5,821,337, the relevant text of which is incorporated herein by reference.  
20       Humanized antibodies may be designed against any surface-expressed antigenic protein or peptide, including but not limited to PSMA, PSCA, caveolin, POV1, HER2/neu and p27KIP1. Monoclonal antibodies targeted to such proteins or peptides may be generated by techniques well known in the art, as summarized above. Modification of such antibodies for optimal expression in human subjects, without provoking an anti-mouse immune response, may be accomplished by the  
25       following methods.

          Such humanization could be carried out, for example, by joining the variable segments of the genes from a mouse monoclonal antibody with human constant segments, such as gamma 1 and gamma 3 (see, *e.g.*, Liu *et al.*, 1987). A typical humanized therapeutic chimeric antibody is thus a hybrid protein composed of the variable or antigen-binding domain from a  
30       non-human (*e.g.*, mouse) antibody and the constant or effector domain from a human antibody. Preferably, the entire rodent variable region is not retained, as in some cases this may lead to a human anti-mouse immune response.



An immunoglobulin variable region comprises light and heavy chains having a "framework" region interrupted by hypervariable regions, called complementarity determining regions (CDR's). The CDRs comprise amino acid sequences which together define the binding affinity and specificity of the natural variable region of a native immunoglobulin binding site, or a synthetic polypeptide which mimics this function. CDRs are not necessarily wholly homologous to hypervariable regions of natural immunoglobulin molecules, and also may include specific amino acids or amino acid sequences which flank the hypervariable region and have heretofore been considered framework not directly determinative of complementarity. The framework regions are found naturally occurring between CDRs in immunoglobulins. These sequences may be derived in whole or part from the same immunoglobulin as the CDRs, or in whole or part from a different immunoglobulin. For example, in order to enhance biocompatibility of a humanized antibody to be administered to a human, the framework sequences can be derived from a human immunoglobulin so that the resulting humanized antibody will be less immunogenic than a murine monoclonal antibody.

One form of immunoglobulin (*e.g.*, F(ab')<sub>2</sub>, Fv, Fab, bifunctional antibodies, antibodies, *etc.*) constitutes the basic structural unit of an antibody. This form is a tetramer and consists of two identical pairs of immunoglobulin chains, each pair having one light and one heavy chain. In each pair, the light and heavy chain variable regions are together responsible for binding to an antigen, and the constant regions are responsible for the antibody effector functions. In addition to antibodies, immunoglobulins may exist in a variety of other forms including, for example, Fv, Fab, and F(ab')<sub>2</sub>, as well as bifunctional hybrid antibodies (see *e.g.*, Lanzavecchia *et al.*, 1987)) and in single chains (see Huston *et al.*, 1988; Bird *et al.*, 1988).

#### 4.4.1 Preparation of Humanized Antibodies

As used herein, a humanized antibody comprises an immunoglobulin with a human framework region and one or more CDR's from a non-human (usually a mouse or rat) immunoglobulin. The non-human immunoglobulin providing the CDR's is called the "donor" and the human immunoglobulin providing the framework is called the "acceptor." Constant regions need not be present, but if they are, they should be substantially identical to human immunoglobulin constant regions, *i.e.*, at least about 85-90%, preferably about 95% or more identical. Hence, all parts of a humanized immunoglobulin, except possibly the CDR's, are substantially identical to corresponding parts of natural human immunoglobulin sequences. A humanized antibody thus comprises a humanized light chain and a humanized heavy chain

immunoglobulin. For example, mouse CDR's with or without additional naturally-associated mouse amino acid residues, can be introduced into human framework regions to produce humanized immunoglobulins capable of binding to PSMA, PSCA, caveolin, POV1, HER2/neu or p27KIP1. The humanized immunoglobulins may be utilized alone in substantially pure form, or together with other therapeutic agents.

Framework regions can be identified by homology searches of the GenBank database and then introduced into a particular recombinant immunoglobulin by site-directed mutagenesis to reproduce the corresponding human sequence. Alternatively, homologous human V.sub.H and V.sub.L sequences can be derived from a collection of PCR<sup>TM</sup>-cloned human V regions, after which the human framework sequences can be ligated with murine CDR regions to create humanized V.sub.L and V.sub.H genes. A humanized sFv hybrid thus can be created, for instance, where the human framework regions of the human myeloma antibody are introduced between the murine CDR sequences of a murine monoclonal antibody. The resulting sFv, containing the sequences FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, contains a murine binding site in a human framework.

The extent of the framework region and CDR's have been precisely defined (see, for example, EP 0 239 400, the disclosure of which is specifically incorporated herein by reference). The sequences of the framework regions of different light or heavy chains are relatively conserved within a species. A human framework region is a framework region that is substantially identical (about 85% or more, usually 90-95% or more) to the framework region of a naturally occurring human immunoglobulin. The framework region of an antibody, that is the combined framework regions of the constituent light and heavy chains, serves to position and align the CDR's. The CDR's are primarily responsible for binding to an epitope of an antigen.

Human constant region and rearranged variable region DNA sequences can be isolated in accordance with well known procedures from a variety of human cells, but preferably from immortalized B-cells. Similar methods can be used to isolate non-human immunoglobulin sequences. In particular, by directly sequencing the DNA or RNA in a hybridoma secreting an antibody to a preselected antigen, or by obtaining the sequence from the literature, one skilled in the art can essentially produce any desired CDR and framework sequence. Expressed sequences subsequently may be tested for binding and empirically refined by exchanging selected amino acids in relatively conserved regions, based on observations of trends of amino acid sequences in data bases and/or by using computer-assisted modeling techniques.

Significant flexibility in V.sub.H and V.sub.L design is possible because alterations in amino acid sequences may be made at the DNA level.

In addition to naturally-occurring forms of immunoglobulin chains, modified heavy and light chains can be designed and manufactured utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally-occurring sequence at the primary structure level by several amino acid substitutions, terminal and intermediate additions and deletions, and the like. Alternatively, polypeptide fragments comprising only a portion of the primary structure may be produced, which fragments possess one or more immunoglobulin activities (*e.g.*, binding activity). In particular, it is noted that, like many genes, the immunoglobulin-related genes contain separate functional regions, each having one or more distinct biological activities. In general, modifications of the genes encoding the desired epitope binding components may be readily accomplished by a variety of well-known techniques, such as site-directed mutagenesis (see, *e.g.*, Gillman and Smith, 1979; and Roberts *et al.*, 1987, both of which are incorporated herein by reference).

#### 4.4.2 Expression of Humanized Antibodies

Resulting humanized immunoglobulin DNA constructs are integrated into expression vectors and transfected into appropriate host cells for protein expression. After being produced by cells, the protein may be purified from the cells themselves or recovered from the culture medium. Alternatively, the expression vector may be transfected into cells *in situ* for expression in an individual with prostate cancer.

Insertion of DNAs encoding heavy and light chain variable regions into a vector is easily accomplished when the termini of both the DNAs and the vector comprise compatible restriction sites. In other situations, it may be necessary to modify the termini of the DNAs and/or vector by digesting back single-stranded DNA overhangs generated by restriction endonuclease cleavage to produce blunt ends, or to achieve the same result by filling in the single-stranded termini with an appropriate DNA polymerase. Alternatively, any desired restriction site may be produced by ligating nucleotide sequences (linkers) onto the termini. Such linkers may comprise specific oligonucleotide sequences that define desired restriction sites. The cleaved vector and the DNA fragments may also be modified if required by homopolymeric tailing.

Sequences of the heavy and light chain variable regions of the rodent or other non-human monoclonal antibody to be humanized are optimally aligned and compared with known

human antibody heavy and light chain variable region sequences. Once the sequences are compared, residue identities are noted and percent identities may be determined. All other factors being equal, it is desirable to select a human antibody that has the highest percent identity with the animal antibody. The known human antibody chain sequences are then  
5 evaluated for the presence of unidentified residues and/or ambiguities, which are sequence uncertainties. The most common of such uncertainties are the mistaken identification of an acidic amino acid for an amide amino acid due to loss of ammonia during the sequencing procedure, *e.g.*, incorrect identification of a glutamic acid residue, when the residue actually present in the protein was a glutamine residue. Uncertainties are identified by examination of  
10 an appropriate database. All other factors being equal, it is desirable to select a human antibody chain having as few such ambiguities as possible.

Antibody chain variable regions contain intra-domain disulfide bridges. The distance (number of residues) between the cysteine residues comprising these bridges is referred to as the Pin-region spacing (see, *e.g.*, Chothia *et al.*, 1987). It is desirable that the Pin-region  
15 spacing of a human antibody selected be similar or identical to that of the animal antibody. It is also desirable that the human sequence Pin-region spacing be similar to that of a known antibody 3-dimensional structure, to facilitate computer modeling. Based upon the foregoing criteria, the human antibody having the best overall combination of desirable characteristics may be used as the framework for humanization of the rodent antibody. A determination of  
20 which rodent or other non-human antibody variable region sequence should be used for insertion into the human framework may also be made. Such considerations may be made based on whether variable region residues comprise CDR structural loops.

#### 4.4.3 Single Chain Binding Proteins

25 The foregoing discussion and sequences disclosed herein can be used to produce single-chain binding proteins against PSMA, PSCA, caveolin, POV1, HER2/neu or p27KIP1. Such single-chain binding proteins comprise linked heavy and light chain fragments of the Fv region, or biosynthetic antibody binding sites (BABS) (see, *e.g.*, Bird *et al.*, 1988; and Huston *et al.*, 1988). Single-domain antibodies comprising isolated heavy-chain variable domains can also be  
30 prepared (Ward *et al.*, 1989).

Two or more CDRs can also be coupled together in a polypeptide, either directly or by a linker sequence. One or more of the CDRs can also be engineered into another (non-

immunoglobulin) polypeptide or protein, thereby conferring IL-4 binding capability on the polypeptide or protein.

DNAs which encode the heavy and light chain variable regions of an antibody or the CDRs therefrom can be prepared by standard methods using the nucleic acid sequence information provided herein. For example, such DNA can be chemically synthesized (Matteucci *et al.*, 1981; Yoo *et al.*, 1989). Alternatively, since the sequence of the gene and the site specificities of the many available restriction endonucleases are known, one skilled in the art can readily identify and isolate the gene from the genomic DNA of an appropriate hybridoma and cleave the DNA to obtain the desired sequences. PCR™ could be used to obtain the same result (Daugherty *et al.*, 1991). Primers used for PCR™ can, if desired, be designed to introduce appropriate new restriction sites. to facilitate incorporation into a given vector.

Still another method for obtaining DNAs encoding the heavy and light chain variable regions of specific monoclonal antibodies entails the preparation of cDNA, using mRNA isolated from an appropriate hybridoma as a template. Variable regions may be cloned from the cDNA using standard methods (see, *e.g.*, Wall *et al.*, 1978; Zalsut *et al.*, 1980; Cabilly *et al.*, 1984; Boss *et al.*, 1984; Amster *et al.*, 1980; and U.S. Pat. No. 4,642,234). It may also be advantageous to make more substantial modifications. For example, Roberts *et al.* (1987) have produced an antibody with enhanced affinity and specificity by removing two charged residues at the periphery of the combining site by site-directed mutagenesis.

Due to the degeneracy of the genetic code, many different nucleotide sequences can encode recombinant humanized antibodies and the CDRs therein. The codons can be selected for optimal expression in prokaryotic or eukaryotic systems. Such functional equivalents are also a part of this invention. Moreover, those skilled in the art are aware that there can be conservatively modified variants of polypeptides and proteins in which there are minor amino acid substitutions, additions or deletions that do not substantially alter biological function. Thus, it is well within the skill of the art, *e.g.*, by chemical synthesis or by the use of modified PCR™ primers or site-directed mutagenesis, to modify the DNAs of this invention to make such variants if desired.

30

#### 4.5 Immunodetection Assays

In still further embodiments, the present invention concerns immunodetection methods for binding, purifying, removing, quantifying or otherwise generally detecting biological components.

The encoded proteins or peptides of the present invention may be employed to detect antibodies having reactivity therewith, or, alternatively, antibodies prepared in accordance with the present invention, may be employed to detect the encoded proteins or peptides. An embodiment of the present invention comprises assay of the expression of reporter genes linked to native or modified prostate specific promoters, such as deletion mutants of such promoters. Another embodiment involves assay of the expression of reporter genes linked to prostate specific promoters in the presence of putative activators or inhibitors of the promoter. The steps of various useful immunodetection methods have been described in the scientific literature (Nakamura *et al.*, 1987a; Nakamura *et al.*, 1987b ).

10           In general, immunobinding methods include obtaining a sample suspected of containing a protein, peptide or antibody, and contacting the sample with an antibody or protein or peptide in accordance with the present invention, as the case may be, under conditions effective to allow the formation of immunocomplexes.

          The immunobinding methods include methods for detecting or quantifying the amount of a reactive component in a sample, which methods require the detection or quantitation of any immune complexes formed during the binding process. One would obtain a sample suspected of containing a prostate disease-marker encoded protein, peptide or a corresponding antibody, and contact the sample with an antibody or encoded protein or peptide, as the case may be, and then detect or quantify the amount of immune complexes formed under the specific conditions. One may alternatively obtain an extract from cells containing a reporter gene operably linked to a promoter and detecting or quantifying the amount of immune complexes formed, using an antibody against the protein product of the reporter gene.

          In terms of antigen detection, the biological sample analyzed may be any sample, such as a prostate or lymph node tissue section or specimen, a homogenized tissue extract, an isolated cell, a cell membrane preparation, a blood lymphocyte separated or purified forms of any of the above protein-containing compositions, or even any biological fluid, including blood, lymphatic fluid, or seminal fluid. Culture of cells under *in vitro* conditions may provide other types of samples for analysis. Assay of reporter genes expressing a secreted protein product may be accomplished by quantifying the protein present in a cell culture supernatant.

30           Contacting the chosen sample with the protein, peptide or antibody under conditions effective and for a period of time sufficient to allow the formation of immune complexes (primary immune complexes) is generally a matter of simply adding the composition to the sample and incubating the mixture for a period of time long enough for the antibodies to form immune

complexes with, *i.e.*, to bind to, any antigens present. After this time, the sample-antibody composition, such as a tissue section, ELISA plate, dot blot or Western blot, will generally be washed to remove any non-specifically bound antibody species, allowing only those antibodies specifically bound within the primary immune complexes to be detected.

5       The detection of immunocomplex formation is well known in the art and may be achieved through the application of numerous approaches. These methods are generally based upon the detection of a label or marker, such as any radioactive, fluorescent, biological or enzymatic tags or labels of standard use in the art. U.S. Patents concerning the use of such labels include Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241, each  
10       incorporated herein by reference. Of course, one may find additional advantages through the use of a secondary binding ligand such as a second antibody or a biotin/avidin ligand binding arrangement, as is known in the art.

Further methods include the detection of primary immune complexes by a two step approach. A second binding ligand, such as an antibody, that has binding affinity for the encoded  
15       protein, peptide or corresponding antibody is used to form secondary immune complexes, as described above. After washing, the secondary immune complexes are contacted with a third binding ligand or antibody that has binding affinity for the second antibody, again under conditions effective and for a period of time sufficient to allow the formation of immune complexes (tertiary immune complexes). The third ligand or antibody is linked to a detectable  
20       label, allowing detection of the tertiary immune complexes thus formed. This system may provide for signal amplification if this is desired.

#### 4.6   Detection and Quantitation of RNA Species

Gene expression may also be determined by analysis of the RNA products of a reporter  
25       gene or other expression construct. For assay of low level transcripts, quantitative analysis may involve an initial amplification of the gene product. Nucleic acid used as a template for amplification is isolated from cells according to standard methodologies. (Sambrook *et al.*, 1989) Where RNA is used, it may be desired to convert the RNA to a complementary cDNA. In one embodiment, the RNA is whole cell RNA and is used directly as the template for amplification.

30       Pairs of primers that selectively hybridize to nucleic acids corresponding to the gene of interest are contacted with the isolated nucleic acid under conditions that permit selective hybridization. The term primer, as defined herein, is meant to encompass any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process.

Typically, primers are oligonucleotides from ten to twenty base pairs in length, but longer sequences may be employed. Primers may be provided in double-stranded or single-stranded form, although the single-stranded form is preferred. Once hybridized, the nucleic acid:primer complex is contacted with one or more enzymes that facilitate template-dependent nucleic acid synthesis. Multiple rounds of amplification, also referred to as "cycles," are conducted until a sufficient amount of amplification product is produced.

Next, the amplification product is detected. In certain applications, the detection may be performed by visual means. Alternatively, the detection may involve indirect identification of the product *via* chemiluminescence, radioactive scintigraphy of incorporated radiolabel or fluorescent label or even *via* a system using electrical or thermal impulse signals (Affymax technology; Bellus, 1994).

#### 4.6.1 Template Dependent Amplification Methods

A number of template dependent processes are available to amplify the oligonucleotide sequences present in a given template sample. One of the best known amplification methods is the polymerase chain reaction (referred to as PCR<sup>TM</sup>) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*, 1990, each of which is incorporated herein by reference in its entirety.

A reverse transcriptase PCR<sup>TM</sup> amplification procedure may be performed in order to quantify the amount of mRNA amplified. Methods of reverse transcribing RNA into cDNA are well known and described in Sambrook *et al.*, 1989. Alternative methods for reverse transcription utilize thermostable DNA polymerases. These methods are described in WO 90/07641 filed December 21, 1990. Polymerase chain reaction methodologies are well known in the art. The most preferred methods of RT-PCR<sup>TM</sup> are as described in US Application Serial No. 08/692,787, which is incorporated herein by reference in its entirety.

Another method for amplification is the ligase chain reaction ("LCR"), disclosed in EPO 320,308, incorporated herein by reference in its entirety. U.S. Patent 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence. Qbeta Replicase, described in PCT Application No. PCT/US87/00880, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA which has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence which may then be detected.



An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[alpha-thio]-triphosphates in one strand of a restriction site may also be useful in the amplification of nucleic acids in the present invention (Walker *et al.*, 1992). Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.*, nick translation.

Still other amplification methods described in GB Application No. 2,202,328, and in PCT Application No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR (Kwoh *et al.*, 1989; Gingeras *et al.*, PCT Application WO 88/10315, incorporated herein by reference in their entirety). Davey *et al.*, EPO 329,822 (incorporated herein by reference in its entirety) disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention.

Miller *et al.*, PCT Application WO 89/06700 (incorporated herein by reference in its entirety) disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter region/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic, *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "race" and "one-sided PCR<sup>TM</sup>" (Frohman, 1990; Ohara *et al.*, 1989).

#### 4.6.2 Separation Methods

Following amplification, it may be desirable to separate the amplification product from the template and the excess primer for the purpose of determining whether specific amplification has occurred. In one embodiment, amplification products are separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods. See Sambrook *et al.*, 1989.

Alternatively, chromatographic techniques may be employed to effect separation. There are many kinds of chromatography which may be used in the present invention: adsorption, partition, ion-exchange and molecular sieve, and many specialized techniques for using them including column, paper, thin-layer and gas chromatography (Freifelder, 1982).

#### 4.6.3 Identification Methods

Amplification products must be visualized in order to confirm amplification. One typical visualization method involves staining of a gel with ethidium bromide and visualization under UV  
5 light. Alternatively, if the amplification products are integrally labeled with radio- or fluorometrically-labeled nucleotides, the amplification products may then be exposed to x-ray film or visualized under the appropriate stimulating spectra, following separation.

In one embodiment, visualization is achieved indirectly. Following separation of amplification products, a labeled, nucleic acid probe is brought into contact with the amplified  
10 marker sequence. The probe preferably is conjugated to a chromophore but may be radiolabeled. In another embodiment, the probe is conjugated to a binding partner, such as an antibody or biotin, where the other member of the binding pair carries a detectable moiety.

In one embodiment, detection is by Southern blotting and hybridization with a labeled probe. The techniques involved in Southern blotting are well known to those of skill in the art and  
15 may be found in many standard books on molecular protocols. See Sambrook *et al.*, 1989. Briefly, amplification products are separated by gel electrophoresis. The gel is then contacted with a membrane, such as nitrocellulose, permitting transfer of the nucleic acid and non-covalent binding. Subsequently, the membrane is incubated with a chromophore-conjugated probe that is capable of hybridizing with a target amplification product. Detection is by exposure of the  
20 membrane to x-ray film or ion-emitting detection devices.

One example of the foregoing is described in U.S. Patent No. 5,279,721, incorporated by reference herein, which discloses an apparatus and method for the automated electrophoresis and transfer of nucleic acids. The apparatus permits electrophoresis and blotting without external manipulation of the gel and is ideally suited to carrying out methods according to the present  
25 invention.

#### 4.6.4 Other Assays

Methods for detecting mutations in genomic DNA, cDNA or RNA samples may be employed, depending on the specific situation. A number of different methods have been used to detect point mutations, including denaturing gradient gel electrophoresis ("DGGE"),  
5 restriction enzyme polymorphism analysis, chemical and enzymatic cleavage methods, and others. The more common procedures currently in use include direct sequencing of target regions amplified by PCR<sup>TM</sup> (see above) and single-strand conformation polymorphism analysis ("SSCP").

Another method of screening for point mutations is based on RNase cleavage of base  
10 pair mismatches in RNA/DNA and RNA/RNA heteroduplexes. As used herein, the term "mismatch" is defined as a region of one or more unpaired or mispaired nucleotides in a double-stranded RNA/RNA, RNA/DNA or DNA/DNA molecule. This definition thus includes mismatches due to insertion/deletion mutations, as well as single and multiple base point mutations.

15 U.S. Patent No. 4,946,773 describes an RNase A mismatch cleavage assay that involves annealing single-stranded DNA or RNA test samples to an RNA probe, and subsequent treatment of the nucleic acid duplexes with RNase A. After the RNase cleavage reaction, the RNase is inactivated by proteolytic digestion and organic extraction, and the cleavage products are denatured by heating and analyzed by electrophoresis on denaturing polyacrylamide gels.  
20 For the detection of mismatches, the single-stranded products of the RNase A treatment, electrophoretically separated according to size, are compared to similarly treated control duplexes. Samples containing smaller fragments (cleavage products) not seen in the control duplex are scored as positive.

Currently available RNase mismatch cleavage assays, including those performed  
25 according to U.S. Patent No. 4,946,773, require the use of radiolabeled RNA probes. Myers and Maniatis in U.S. Patent No. 4,946,773 describe the detection of base pair mismatches using RNase A. Other investigators have described the use of *E. coli* enzyme, RNase I, in mismatch assays. Because it has broader cleavage specificity than RNase A, RNase I would be a desirable enzyme to employ in the detection of base pair mismatches if components can be  
30 found to decrease the extent of non-specific cleavage and increase the frequency of cleavage of mismatches. The use of RNase I for mismatch detection is described in literature from Promega Biotech. Promega markets a kit containing RNase I that is shown in their literature to cleave three out of four known mismatches, provided the enzyme level is sufficiently high.

The RNase Protection assay (RPA) was adapted for detection of single base mutations. In this type of RNase A mismatch cleavage assay, radiolabeled RNA probes transcribed *in vitro* from wild-type sequences, are hybridized to complementary target regions derived from test samples. The test target generally comprises DNA (either genomic DNA or DNA amplified by cloning in plasmids or by PCR<sup>TM</sup>), although RNA targets (endogenous mRNA) have occasionally been used. If single nucleotide (or greater) sequence differences occur between the hybridized probe and target, the resulting disruption in Watson-Crick hydrogen bonding at that position ("mismatch") can be recognized and cleaved in some cases by single-strand specific ribonuclease. To date, RNase A has been used almost exclusively for cleavage of single-base mismatches, although RNase I has recently been shown as useful also for mismatch cleavage. There are recent descriptions of using the MutS protein and other DNA-repair enzymes for detection of single-base mismatches.

#### 4.6.5 Kit Components

All the essential materials and reagents required for detecting specific nucleic acids in a sample may be assembled together in a kit. The kit generally will comprise preselected primer pairs for one or more specific genes. For example a kit may include primers and/or probes for use in any molecular biology assay known to those of skill in the art, such as RT-PCR<sup>TM</sup>, *in situ* hybridization, Northern analysis and/or RPA, to detect RNA markers of normal tissue, BPH tissue, confined tumor tissue or metastatically progressive tumor tissue, or any combination of these. Also included may be enzymes suitable for amplifying nucleic acids including various polymerases (RT, Taq, *etc.*), deoxynucleotides and buffers to provide the necessary reaction mixture for amplification. Preferred kits may also comprise primers for the detection of a control, non-differentially expressed RNA such as  $\beta$ -actin, for example.

The kits generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each marker primer pair. Preferred pairs of primers for amplifying nucleic acids are selected to amplify the sequences designated herein as SEQ ID NO:1, SEQ ID NO:2 or SEQ ID NO:3.

In certain embodiments, kits will comprise hybridization probes specific for differentially expressed genes. The probes are designed to hybridize to a sequence or a complement of a sequence designated herein as SEQ ID NO:1, SEQ ID NO:2 or SEQ ID NO:3. Such kits generally will comprise, in suitable means for close confinement, distinct containers for each individual reagent and enzyme as well as for each marker hybridization probe.

#### 4.7 Identification of Inhibitors and Activators of the Prostate Specific Transglutaminase Promoter and Therapeutic Uses

As stated above, evidence suggests a role for prostate specific transglutaminase, cytokeratin 15, and semenogelin II in prostate cancers. One embodiment of the present invention involves a cell-based assay technique for identifying and evaluating chemical compounds and agents which affect the production of prostate specific transglutaminase, thereby identifying chemotherapeutic compounds for use in the treatment of prostate cancer. This cell-based assay also is believed to work equally well in assessing compounds for their stimulation or inhibition of prostate specific transglutaminase production in prostate cancers.

Specifically, cells are transfected with an expression vector comprising a DNA sequence encoding a promoter region of prostate specific transglutaminase, operatively linked to a reporter gene encoding an assayable product. Examples of reporter genes are well known in the art and include CAT (chloramphenical acetyltransferase), GUS ( $\beta$ -glucuronidase) and luciferase. The cells are then cultured under conditions which permit expression of the assayable product. The prostate specific transglutaminase promoter region is preferably cloned from genomic DNA but may be synthesized *de novo*. A particular example of a promoter region for prostate specific transglutaminase is provided herein as SEQ ID NO:15.

After transfection with the expression vector, the cells are incubated with at least one compound suspected of possessing regulatory activity for prostate specific transglutaminase expression. Chemical agents and factors can be identified by their ability to modulate the expression of the reporter gene and thereby increase or decrease the production of the assayable product. Such chemical compounds are selected from small chemical libraries, peptide libraries, and/or collections of natural products, such as hormones. Examples of compounds that may be assayed include, but are not limited to, hormones such as androgens, glucocorticoids and progesterone.

The present invention is distinguished from other techniques for identifying chemical compounds, as it specifically identifies chemical compounds, agents, factors and other substances which affect the expression of prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II by cells. These agents are identified by their capacity to affect the activity of the prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II promoter regions (for example, SEQ ID NO:14 or SEQ ID NO:15). A change in activity of the promoter regions is measured by a correspondent increase or decrease in production of the

protein encoded by the reporter gene. Production of a reporter gene protein product can be determined by well-known methods. For example, expression of the CAT gene can be quantified by the formation of a radio-labeled acetylated form of chloramphenicol, assayed by thin-layer chromatography followed by liquid scintillation counting or equivalent techniques.

5        Thus, decrease in the production of, for example, luciferase under the control of a prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II promoter region, indicates that the promoter activity is being suppressed by the compound being tested. An increase in the production of luciferase is indicative of stimulation of promoter activity. The effect of a given compound on production of the assaying product is assumed to reflect the  
10 effect on the expression of the prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II genes that would occur in a cell treated with the compound.

Ultimately, when cancer patients are treated with chemical compounds shown to increase the promoter activities of prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II, the production of prostate specific transglutaminase, cytokeratin 15, and/or  
15 semenogelin II proteins by prostate tumor cells will be stimulated. Therefore, compounds identified by this assay technique that increase prostate specific transglutaminase, cytokeratin 15, and semenogelin II promoter activity can be used in the treatment of metastatic prostate cancers, as well as other conditions where a reduction in prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II production is associated with detrimental effects on  
20 patient health.

#### 4.8 Prostate Specific Promoter Regions

One embodiment of the present invention concerns the identification of novel prostate specific promoters. These techniques are generally based upon the "cloning" of a DNA  
25 molecule comprising the promoter from a genomic human DNA library. This can be achieved, for example, by cloning of a genomic DNA molecule containing a prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II promoter. Alternatively, having knowledge of the prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II promoter sequence, the promoter region may be synthesized according to standard techniques.

30        The first step in a cloning procedure is the screening of an appropriate DNA library, such as a human genomic library. The screening procedure may be based on the hybridization of oligonucleotide probes, designed from a consideration of portions of the genes encoding prostate specific transglutaminase, semenogelin II or cytokeratin 15, particularly with respect to

the 5' ends of such genes. The operation of such screening protocols is well known to those of skill in the art and is described in detail in the scientific literature. Nucleotide sequences in accordance with the 5' ends of SEQ ID NO:1, SEQ ID NO:2 or SEQ ID NO:3 may be used as probes, as described in the preceding sections, in screening protocols. These probes would be  
5 hybridized under high stringency conditions with a human genomic library.

Alternatively, the sequences of the prostate specific transglutaminase promoter (SEQ ID NO:15) or semenogelin II promoter (SEQ ID NO:14) may be used to design probes for library screening. In this case, it may be desired to hybridize such probes under low-stringency conditions to DNAs from a human genomic library, in order to identify related promoter  
10 sequences that are potentially differentially expressed in prostate cancers. The present invention provides a prostate specific transglutaminase promoter sequence in SEQ ID NO:15. Additionally, a 4409 bp fragment of the semenogelin II promoter region is set forth in SEQ ID NO:14.

Much of the thinking about how promoters are organized derives from analyses of  
15 several viral promoters, including those for the HSV thymidine kinase (tk) and SV40 early transcription units. These studies have shown that promoters are composed of discrete functional modules or regulatory elements, each consisting of approximately 7-20 bp of DNA, and containing one or more recognition sites for transcriptional activator or repressor proteins. A listing of potential regulatory elements contained within the promoter region of prostate  
20 specific transglutaminase (SEQ ID NO:15) is provided in Example 6 below.

The prostate specific transglutaminase promoter (SEQ ID NO:15) may find wide utility in directing the expression of any gene which one desires to have expressed in prostate tissues or prostate derived tissues, such as prostate cancer. By including a prostate specific transglutaminase promoter region with transformation constructs comprising one or more genes  
25 operatively linked to this promoter region, one may increase or regulate the level of expression of these genes, preferably in prostate derived tissues. Alternatively, the prostate specific transglutaminase promoter may be included in conjunction with any other animal or viral promoters, enhancers, or other regulatory sequences for the enhanced expression of one or more selected genes.

30 At least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase

gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

Additional promoter elements regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the tk promoter, the spacing between promoter elements can be increased to 50 bp or more apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either co-operatively or independently to activate transcription.

In addition to the unmodified prostate specific transglutaminase promoter sequence, given in SEQ ID NO:15, the current invention includes derivatives of this sequence. In particular, the present disclosure provides the teaching for one of skill in the art to make and use derivatives of this sequence. For example, the disclosure provides the teaching for one of skill in the art to delimit the functional elements within the prostate specific transglutaminase promoter sequence and to delete any non-essential elements. Functional elements could also be modified to increase the utility of the sequences of the invention for a particular application. For example, a functional region within the prostate specific transglutaminase promoter of the invention could be modified to cause or increase tissue-specific expression. Such changes could be made by site-specific mutagenesis, for example, as described below.

#### 4.8.1 Derivatives of Prostate Specific Promoter Sequences

One aspect of the invention provides derivatives of human prostate specific promoters. In particular, the current invention includes sequences which have been derived from the human prostate specific transglutaminase promoter region and the semenogelin II promoter region. One means for preparing derivatives of such promoters comprises introducing mutations into the promoter sequences, for example, the sequences given in SEQ ID NO:14 or SEQ ID NO:15. Such mutants may potentially have enhanced, reduced, or altered function relative to the native sequence or alternatively, may be silent with regard to function.

Mutagenesis may be carried out at random and the mutagenized sequences screened for function. Alternatively, particular sequences which provide the prostate specific transglutaminase promoter region with desirable expression characteristics could be identified and these or similar sequences introduced into other related or non-related sequences via



mutation. Examples of sequences within the prostate specific transglutaminase promoter that may serve as preferential targets for mutagenesis are described in Table 8 below. Similarly, non-essential elements may be deleted without significantly altering the function of the elements. It is further contemplated that one could mutagenize these sequences in order to enhance their utility in expressing transgenes, especially in a gene therapy construct in humans.

The means for mutagenizing a DNA segment comprising the prostate specific transglutaminase promoter sequence of the current invention are well-known to those of skill in the art. Mutagenesis may be performed in accordance with any of the techniques known in the art, such as, and not limited to, synthesizing an oligonucleotide having one or more mutations within the sequence of a particular regulatory region. In particular, site-specific mutagenesis is a technique useful in the preparation of promoter mutants, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, by introducing one or more nucleotide sequence changes into the DNA.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to about 75 nucleotides or more in length is preferred, with about 10 to about 25 or more residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art, as exemplified by various publications. As will be appreciated, the technique typically employs a phage vector which exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids also are routinely employed in site directed mutagenesis to eliminate the step of transferring the gene of interest from a plasmid to a phage.

Alternatively, the use of PCR<sup>TM</sup> with commercially available thermostable enzymes such as *Taq* polymerase may be used to incorporate a mutagenic oligonucleotide primer into an amplified DNA fragment that can then be cloned into an appropriate cloning or expression vector. The PCR<sup>TM</sup>-mediated mutagenesis procedures of Tomic *et al.* (1990) and Upender *et al.* (1995) provide two examples of such protocols.

The preparation of sequence variants of the selected promoter or intron-encoding DNA segments using site-directed mutagenesis is provided as a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of DNA sequences may be obtained. For example, recombinant vectors encoding the  
5 desired promoter sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U.S.  
10 Patent No. 4,237,224, incorporated herein by reference. A number of template dependent processes are available to amplify the target sequences of interest present in a sample, such methods being well known in the art and specifically disclosed herein.

One efficient, targeted means for preparing mutagenized promoters or enhancers relies upon the identification of putative regulatory elements within the target sequence. These can  
15 be identified, for example, by comparison with known promoter sequences. Sequences which are shared among genes with similar functions or expression patterns are likely candidates for the binding of transcription factors and are likely elements to confer tissue specific expression patterns. Since prostate specific transglutaminase is directed by its promoter to confer tissue specific expression in prostate derived tissues, comparisons to the promoter regions and genes  
20 of other prostate specific genes, and other genes that are non-exclusively expressed in prostate tissues may be used to identify prostate specific regulatory elements. As used herein, "prostate specific" means expressed in prostate or prostate derived tissues with little, if any, relative expression in other tissues, as detected by techniques such as Northern analysis, relative quantitative reverse transcriptase PCR™, or other quantitative techniques.

25 An example of a prostate specific genes, and genes expressed in prostate tissues whose promoter sequences may be compared to the sequences of the present invention include prostate specific antigen (Zhang *et al.*, 1997; Shan *et al.*, 1997; Cleutjens *et al.*, 1997), the androgen receptor (Dai and Burnstein, 1996); probasin (Yan *et al.*, 1997; Matusik, U.S. Patent No. 5,783,681, July 21, 1998); prostatic acid phosphatase (Shan *et al.*, 1997; Zelevianski *et al.*,  
30 1998); prostate specific glandular kallikrein (Shan *et al.*, 1997); prostate specific membrane antigen (Israeli *et al.*, 1993); testosterone-repressed prostate message-2/clusterin (Wong *et al.*, 1994); the slp gene (Adler *et al.*, 1991; Alder *et al.*, 1993; Scarlett and Robins, 1995), the human glandular kallikrein-1 (Young *et al.*, 1995); the human renal/pancreatic kallikrein

(Young *et al.*, 1995); seminal plasma protein (Mbikay *et al.*, 1987); homeobox gene NKX3.1 (Prescott *et al.*, 1998) and prostate stem cell antigen (Reiter *et al.*, 1998).

One of skill in the art will recognize that regulatory elements may be included in regions of the gene other than the 5'-untranslated region, and comparison of coding and  
5 3'-noncoding regions of prostate expressed genes to the prostate specific transglutaminase gene, including the coding and non-coding regions, may identify putative regulatory elements. Confirmation of putative regulatory elements can be achieved by deletion, duplication, or other alteration or mutation of each putative regulatory region followed by functional analysis of each construct by assay of a reporter gene which is functionally attached to each construct. As  
10 such, once a starting promoter sequence is provided, any of a number of different functional mutants of the starting sequence could be readily prepared using methods well known in the art (Zhang *et al.*, 1997).

Mutation, alteration, duplication, or truncation mutants of the prostate specific transglutaminase promoter region of the invention could be randomly prepared or prepared by  
15 selection of regions identified as containing putative regulatory elements, and then assayed. With this strategy, a series of constructs are prepared, each containing a different portion of the clone (a subclone), either mutated or altered or containing wild-type sequence, and these constructs are then screened for activity. A suitable means for screening for activity is to attach such a promoter construct to a selectable or screenable marker gene and to assay for gene  
20 expression. Preferred cells to conduct such assays would be prostate or prostate derived cells, androgen-repressed human prostate cancer cell lines such as ARCaP, androgen-dependent cell lines such as LNCaP and androgen-independent cell lines such as C4-2, PC3 and DU145, though other non-prostate derived cells can be used. Additionally, such assays may be conducted in a transgenic animal produced by incorporation of the construct as a transgene, or  
25 in an animal or organ transfected with the construct as a vector.

Other assays may be used to identify responsive elements in a promoter region or gene. Such assays will be known to those of skill in the art (see for example, Sambrook *et al.*, 1989; Zhang *et al.*, 1997; Shan *et al.*, 1997; Dai and Burnstein, 1996; Cleutjens *et al.*, 1997; Ng *et al.*, 1994; Shida *et al.*, 1993), and include DNase I footprinting studies, Electromobility Shift Assay  
30 patterns (EMSA), the binding pattern of purified transcription factors, effects of specific transcription factor antibodies in inhibiting the binding of a transcription factor to a putative responsive element, Western analysis, nuclear run-on assays, and DNA methylation interference analysis.

Prerferred promoter constructs may be identified that retain the desired, or even enhanced, activity, *i.e.*, prostate tissue or prostate cancer specific regulation and/or expression. The smallest segment required for activity may be identified through comparison of the selected deletion or mutation constructs. Once identified, such segments may be duplicated, mutated, or combined with other known or regulatory elements and assayed for activity or regulatory properties. Promoter region sequences used to identify regulatory elements can also be used to identify and isolate transcription factors that bind a putative regulatory sequence or element, according to standard methods of protein purification, such as affinity chromatography, as discussed above.

These techniques are commonly used by those of skill in the art to identify regulatory elements in a promoter or gene, including tissue specific regulatory elements for prostate specific or expressed genes (Zhang *et al.*, 1997). Such identified or isolated regions or elements of the prostate specific transglutaminase promoter region, whether wild-type or altered, may be used to construct vectors to express operatively linked genes or sequences. Additionally, other promoters, or regulatory elements, including enhancers, may be combined with or operatively linked to sequences of the prostate specific transglutaminase promoter region in a construct.

Preferred regulatory elements may be from prostate tissue specific or expressed genes. Examples of elements identified in the promoters of prostate expressed genes that may be combined with the sequences of the present invention include the androgen responsive element (ARE) (Zhang *et al.*, 1997; Dai and Burnstein, 1996), the A and B motifs (Zhang *et al.*, 1997), hMT-IIA-MREa (Zhang *et al.*, 1997), OTF (Zhang *et al.*, 1997), CACCC-box (Zhang *et al.*, 1997), steroid response elements (SRE) (Shan *et al.*, 1997), the hormone response element (HRE) (Beato, M., 1989), and the androgen receptor-binding region (ARBR-1) (Dai and Burnstein, 1996),

Preferrably, identified prostate specific transglutaminase promoter region sequences, whether used alone or combined with additional promoters, enhancers, or regulatory elements, will be induced and/or regulated by an external agent, such as a hormone, transcription factor, enzyme, or pharmaceutical agent, to express operatively linked genes or sequences (Zhang *et al.*, 1997; Shan *et al.*, 1997). Preferred hormones that are contemplated to regulate prostate specific transglutaminase promoter region sequences include androgens, estrogen, progesterone, glucocorticoids, and testosterone. For example, prostate tissue specific promoter constructs may be designed to be regulated by androgen by the inclusion of elements such as

the androgen responsive element (ARE) (Zhang *et al.*, 1997) or the A and B motifs (Zhang *et al.*, 1997). Alternatively, such a construct may be designed to cease expression upon exposure to an external agent. Preferably, constructs of the present invention will be regulated specifically in prostate tissues or prostate derived tissues.

5        Additionally, deletion mutants may be produced and assayed essentially according to Matusik (U.S. Patent No. 5,783,681, July 21, 1998). Plasmids may be constructed containing the prostate specific transglutaminase promoter (SEQ ID NO:15) adjacent to a reporter gene, for example CAT. The construct may be designed to contain additional regulatory sequences, such as polyadenylation, termination and cleavage signals. Deletion mutants may be prepared  
10        by a time course treatment of the isolated prostate specific transglutaminase promoter with Bal 31 exonuclease (for 15, 30, 45, 60 and 75 sec, for example). Following limited digestion, the promoter sequence may be ligated to appropriate linker sequences and reinserted into the CAT expression vector. After transformation into an appropriate host cell, such as *E. coli*, clones containing deletion mutants may be screened and their plasmid DNAs digested with restriction  
15        enzymes that cut at sites flanking the promoter sequence. Promoter size may be determined by agarose gel electrophoresis according to standard techniques.

Following selection of a range of deletion mutants of varying size, the activities of the deleted promoters for expression of the linked CAT gene may be determined according to standard protocols. For this purpose, it would be desirable to assay promoter activity in cells of  
20        human prostate origin, such as PC-3, LNCaP, DU145, C4-2, C4-2Ln and C4-2B cell line 2B (Chung *et al.*, 1994, *Cancer Research*, 54:2577-2581.)

The precise nature of the deleted portion of the promoter may be determined using standard DNA sequencing, such as Sanger dideoxy termination sequencing, to identify which promoter sequences have been removed in each of the assayed deletion mutants. Thus, a  
25        correlation may be obtained between the presence or absence of specific elements within the prostate specific transglutaminase promoter sequence and changes in activity of the linked reporter gene.

#### 4.9 Transformation and Expression Constructs

30        The construction of vectors which may be employed in conjunction with tissue culture or animal or human transformation techniques will be known to those of skill of the art in light of the present disclosure (see for example, Sambrook *et al.*, 1989; Gelvin *et al.*, 1990). The techniques of the current invention are not limited to any particular DNA sequences in

conjunction with the prostate specific transglutaminase promoter region or other prostate specific promoter regions identified by the methods of the invention.

Throughout this application, the term "expression construct" is meant to include any type of genetic construct containing a nucleic acid coding for a gene product in which part or  
5 all of the nucleic acid sequence is capable of being transcribed. The transcript may be translated into a protein, but it need not be. In certain embodiments, expression includes both transcription of a gene and translation of mRNA into a protein. In other embodiments, expression only includes transcription of the nucleic acid into an RNA species, without subsequent translation.

10 The expression constructs, commonly referred to as vectors, that can be utilized in a cell-based assay of promoter activity may vary considerably. An expression vector will comprise at least a promoter region, for example, the prostate specific transglutaminase promoter region (SEQ ID NO:15). The vectors may be "standard" expression vectors, *i.e.*, plasmids that contain one or more effector genes and regulatory elements required for  
15 expression of the effector gene in cells. Expression vectors include any plasmid, cosmid or phage construct that is capable of supporting expression of encoded genes in mammalian cells, such as the pUC or Bluescript<sup>TM</sup> plasmid series. Alternatively, these vectors may be more complex, such as the viral vectors discussed below. Vectors may also include structures that assist in replication, such as origins of replication. In addition, almost all expression vectors  
20 contain multipurpose cloning regions. Finally, expression vectors typically have selectable markers, often in the form of antibiotic resistance genes, that permit selection of cells that carry these vectors.

One use of the sequences of the present invention will be in directing the expression of  
a selected gene which encodes a particular protein or polypeptide product. However, the  
25 selected genes also may be DNA segments that do not encode proteins, exemplified by DNAs encoding anti-sense RNA molecules or ribozymes. It is contemplated that where an expressible gene that is not a marker gene is employed in combination with a marker gene, one may employ the separate genes on either the same or different DNA segments for transformation. In the latter case, the different vectors are delivered concurrently to recipient cells to maximize  
30 cotransformation.

The choice of the particular selected genes used in accordance with the prostate specific transglutaminase promoter region for transformation of recipient cells will depend on the purpose of the transformation. One of the purposes of transformation of prostate tissue or

prostate derived tissue would be to introduce a gene that would be toxic to, or inhibit the growth of prostate cancer cells. Additionally, a gene may be introduced to affect the regulation of other genes, such as oncogenes, or to enhance an immune response to prostate tissues or prostate derived tissues. Such a gene may also be toxic or detrimental to normal prostate tissue or other tissues, but the benefit of killing or inhibiting the growth of prostate cancer would outweigh the loss of healthy prostate tissue.

In certain embodiments, it is contemplated that a recipient cell will be transformed with a transformation construct. Two or more transgenes can be created in a single transformation event using either distinct vectors containing different genes or using a single vector incorporating two or more genes. Any two or more transgenes of any description, such as those conferring, for example, a toxic, antigenic or an immunogenic property or quality may be employed as desired.

Such genes may convey a regulatory property to control expression or function of a gene or group of genes, as would be conveyed by an expressed transcription factor. Alternatively, such genes may confer another type of regulatory quality, such as would be produced by an enzymatic inhibitor or allosteric regulator, or serve as a marker gene, or otherwise be useful in diagnosis or therapeutic treatments of prostate diseases. In some embodiments, it is contemplated that one may wish to employ replication-competent viral vectors for tissue culture or animal transformation. However, replication-incompetent viral vectors will be preferred in transfection of an animal or human.

Vectors used for tissue culture or animal transformation may include, for example, plasmids, cosmids, YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes) or any other suitable cloning system. It is contemplated that utilization of cloning systems with large insert capacities will allow introduction of large DNA sequences comprising more than one selected gene. Introduction of such sequences may be facilitated by use of bacterial or yeast artificial chromosomes (BACs or YACs, respectively).

Particularly useful for transformation are expression cassettes that have been isolated from such vectors. DNA segments used for transforming tissue culture or animal cells will generally comprise the cDNA, gene or genes which one desires to introduce into and express in the host cells. These DNA segments can include, in addition to a prostate specific transglutaminase promoter region, structures such as promoters, enhancers, polylinkers, or even regulatory genes. The DNA segment or gene chosen for cellular introduction will often encode a protein which will be expressed in the resultant recombinant cells, resulting in a screenable or

selectable trait, or which will impart an improved phenotype to the resulting transfected or transgenic cell line or animal. However, this may not always be the case. Preferred components likely to be included with vectors used in the current invention are as follows.

#### 5 4.9.1 Regulatory Elements

Constructs prepared in accordance with the current invention will include a prostate specific transglutaminase promoter region or other promoter isolated by the methods disclosed herein. These promoters may be used in the preparation of transformation constructs which comprise a wide variety of other elements. By including the prostate specific transglutaminase  
10 promoter region in a transformation construct, enhanced tissue-specific expression of selected genes may be achieved. Additionally, one may wish to operably link the prostate specific transglutaminase promoter region to one or more enhancer elements in addition to or other than the enhancers associated with prostate specific transglutaminase gene.

One such element that could be used in conjunction with the prostate specific  
15 transglutaminase promoter region is the DNA sequence between the transcription initiation site and the start of the coding sequence of this gene, termed the untranslated leader sequence. The leader sequence can influence gene expression. Compilations of leader sequences have been made to predict optimum or sub-optimum sequences and generate "consensus" and preferred leader sequences (Joshi, 1987). Preferred leader sequences are contemplated to include those  
20 which have sequences predicted to direct optimum expression of the attached gene, *i.e.*, to include a preferred consensus leader sequence which may increase or maintain mRNA stability and prevent inappropriate initiation of translation. The choice of such sequences will be known to those of skill in the art in light of the present disclosure. Sequences that are derived from genes that are highly expressed in tissue culture or animals, and in humans in particular, will be  
25 most preferred.

The selection of enhancers or other regulatory regions or elements for use with the prostate specific transglutaminase promoter may be made based upon the promoter's ability to direct the transfected or transgenic tissue culture cell line or animal's transcriptional activity to the coding region. Useful regulatory elements may be derived from other promoters, including  
30 those that are inducible, viral, synthetic or constitutive (Poszkowski *et al.*, 1989; Odell *et al.*, 1985), temporally regulated, spatially regulated, and spatio-temporally regulated (Chau *et al.*, 1989). Exemplary constitutive promoters include the CaMV 35S promoter (Odell *et al.*, 1985), histone, CaMV 19S (Lawton *et al.*, 1987), *nos* (Ebert *et al.*, 1987), *Adh* (Walker *et al.*, 1987),



sucrose synthase (Yang and Russell, 1990),  $\alpha$ -tubulin, (Wang *et al.*, 1992), cab (Sullivan *et al.*, 1989), PEPCase (Hudspeth and Grula, 1989) and R gene complex-associated promoters (Chandler *et al.*, 1989).

Exemplary tissue-specific promoters include the promoters for prostate specific antigen  
 5 (Takehiko *et al.*, 1998; Gotoh *et al.*, 1998), the androgen receptor (Tilley *et al.*, 1990), probasin  
 (Yan *et al.*, 1997; Matusik, U.S. Patent No. 5,783,681, July 21, 1998), prostatic acid  
 phosphatase (Zelivianski *et al.*, 1998), prostate specific glandular kallikrein (Murtha *et al.*,  
 1993; Shan *et al.*, 1997), prostate specific membrane antigen (Israeli *et al.*, 1993; 1997),  
 testosterone-repressed prostate message-2/clusterin (Wong *et al.*, 1994), relaxin H2 (Brooks *et*  
 10 *al.*, 1998), cytokeratin 15, prostate specific transglutaminase and the semenogelin II promoter  
 regions.

Table 2 lists enhancer elements that may be employed to construct tissue specific  
 promoters regulating the expression of the genes of interest. This list is not intended to be  
 exhaustive of all the possible elements involved in the promotion of gene expression, but  
 15 merely to be exemplary thereof.

Enhancers were originally detected as genetic elements that increased transcription from  
 a promoter located at a distant position on the same molecule of DNA. The basic distinction  
 between enhancers and promoters is operational. An enhancer region as a whole must be able  
 to stimulate transcription at a distance. This need not be true of a promoter region or its  
 20 component elements. Promoters and enhancers are often overlapping and contiguous, often  
 seeming to have a very similar modular organization.

**TABLE 2. Enhancers of Potential Use for Construction of Expression Vectors**

Immunoglobulin Heavy Chain
Immunoglobulin Light Chain
T-Cell Receptor
HLA DQ a and DQ b
b-Interferon
Interleukin-2
Interleukin-2 Receptor

**Table 2 – Continued**

Gibbon Ape Leukemia Virus
MHC Class II 5 or HLA-DRA
b-Actin
Muscle Creatine Kinase
Prealbumin (Transthyretin)
Elastase I
Metallothionein
Collagenase
Albumin Gene
a-Fetoprotein
a-Globin
b-Globin
c-fos
c-HA-ras
Insulin Neural Cell Adhesion Molecule (NCAM)
a1-Antitrypsin
H2B (TH2B) Histone
Mouse or Type I Collagen
Glucose-Regulated Proteins (GRP94 and GRP78)
Rat Growth Hormone
Human Serum Amyloid A (SAA)
Troponin I (TN I)
Platelet-Derived Growth Factor
Duchenne Muscular Dystrophy
SV40 or CMV
Polyoma
Retroviruses
Papilloma Virus
Hepatitis B Virus
Human Immunodeficiency Virus

Transcription enhancers could be used to increase expression. These enhancers are often found 5' to the start of transcription in a promoter that functions in eukaryotic cells, but can often be inserted in the forward or reverse orientation 5' or 3' to the coding sequence. In some instances these 5' enhancing elements are introns. Examples of enhancers include  
5 elements from the CaMV 35S promoter, octopine synthase genes (Ellis *et al.*, 1987), the maize alcohol dehydrogenase gene, the maize shrunk 1 gene; the enhancer identified in the prostate specific antigen promoter (Cleutjens *et al.*, 1997), and promoters from non-plant eukaryotes (*e.g.* yeast; Ma *et al.*, 1988).

The enhancer identified in the prostate specific antigen promoter contains an androgen  
10 receptor binding site. This enhancer was shown to improve transcriptional activity of the PSA proximal promoter in LNCaP cells (Cleutjens *et al.*, 1997) and is contemplated as being useful in improving the expression of constructs of the present invention in prostate tissue or prostate derived tissue.

It also is contemplated that tissue specific expression may be functionally accomplished  
15 by introducing a constitutively expressed gene in combination with an antisense gene that is expressed only in those tissues where the gene product is not desired. For example, a gene operatively linked to a  $\beta$ -actin promoter may be co-transfected or co-transformed with an antisense form of the same gene, operably linked to a human prostate specific transglutaminase promoter sequence. This would result in gene expression in tissues other than prostate. This  
20 may be desirable, for example, to promote the survival of non-prostate cells exposed to a given toxic agent, thereby providing for selective kill of prostate-derived cells. Examples of detoxification genes are well-known to those of skill in the art.

A useful method for identifying tissue-specific promoters is differential display (see, *e.g.*, U.S. Patent No. 5,599,672, the disclosure of which is incorporated herein by reference in  
25 its entirety). In differential display, mRNAs are compared from different tissue types. By identifying mRNA species that are present only in a particular tissue type, or set of tissues types, one can identify the corresponding genes as expressed in a tissue specific manner. The RNAs can be transcribed by reverse transcriptase to produce a cDNA, and the cDNA can be used to isolate clones containing the full-length genes. As disclosed herein, the cDNA can also  
30 be used to isolate promoters, enhancers or terminators from the respective gene using, for example, suppression PCR™.

Another possible element that may be introduced is a matrix attachment region element (MAR), such as the chicken lysozyme A element (Stief *et al.*, 1989). The MAR can be

positioned around an expressible gene of interest to effect an increase in overall expression of the gene and diminish position dependent effects upon incorporation into the tissue culture cell line or animal genome (Stief *et al.*, 1989; Phi-Van *et al.*, 1990).

It is contemplated that expression of some genes in transfected or transgenic tissue culture cell lines or animals will be desired only under specified conditions. Therefore, for some desired traits, inducible expression of genes in transfected or transgenic tissue culture cell lines or animals will be desired. Inducible expression would be most desired to selectively kill or inhibit prostate cancer cells.

#### 10 4.9.2 Terminators

Transformation constructs will typically include the selected gene along with a 3' terminal DNA sequence that acts as a signal to terminate transcription and allow for the polyadenylation of the resultant mRNA. The nature of the polyadenylation signal is not believed to be crucial to the successful practice of the invention, and any such sequence may be employed. Examples include SV40 and bovine growth hormone poly-A sites. Also contemplated as an element of the expression construct is a terminator. These elements can serve to enhance message levels and to minimize read through from the construct into other sequences.

#### 4.9.3 Transit or Signal Peptides

Sequences that are joined to the coding sequence of an expressed gene and removed post-translationally and which facilitate the transport of the protein into or through intracellular or extracellular membranes are termed transit or signal sequences. By facilitating the transport of proteins into compartments inside and outside the cell, these sequences may increase the accumulation of gene product by protecting it from proteolytic degradation. These sequences also allow for additional mRNA sequences from highly expressed genes to be attached to the coding sequence of the genes of interest. Since mRNA being translated by ribosomes is more stable than naked mRNA, the presence of translatable mRNA in front of the gene may increase the overall stability of the mRNA transcript and thereby increase synthesis of the gene product. Since transit and signal sequences are usually post-translationally removed from the initial translation product, the use of these sequences allows for the addition of extra translated sequences that may not appear on the final polypeptide. It is contemplated that targeting of certain proteins may be desirable in order to enhance the stability of the protein (U.S. Patent No. 5,545,818, incorporated herein by reference in its entirety).

Additionally, vectors may be constructed and employed in the intracellular targeting of a specific gene product within the cells of a transfected or transgenic tissue culture cell line or animal or in directing a protein to the extracellular environment. This generally will be achieved by joining a DNA sequence encoding a transit or signal peptide sequence to the coding sequence of a particular gene. The resultant transit, or signal, peptide will transport the protein to a particular intracellular or extracellular destination, respectively, and will then be post-translationally removed.

#### 4.9.4 Marker Genes

One application of the human prostate specific transglutaminase promoter region will be in the expression of marker genes. By employing a selectable or screenable marker gene as the expressible gene of interest, one can provide or enhance the ability to identify transformants. "Marker genes" are genes that impart a distinct phenotype to cells expressing the gene and allow transformed cells to be distinguished from cells that do not have the marker. Such genes may encode either a selectable or screenable marker, depending on whether the marker confers a trait which one can "select" for by chemical means, *i.e.*, through the use of a selective agent (*e.g.*, an antibiotic, such as neomycin, kanamycin, penicillin or other well known agents), or whether it is simply a trait that one can identify through observation or testing, *i.e.*, by "screening" (*e.g.*, GUS, CAT, luciferase). Many examples of suitable marker genes are known to the art and can be employed in the practice of the invention.

Included within the terms selectable or screenable marker genes are genes which encode a "secretable marker" whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers that encode a secretable antigen that can be identified by antibody interaction, or secretable enzymes that can be detected by their catalytic activity. Secretable proteins fall into a number of classes, including small, diffusible proteins detectable, *e.g.*, by ELISA and small enzymes detectable in extracellular solution (*e.g.*,  $\alpha$ -amylase,  $\beta$ -lactamase, phosphinothricin acetyltransferase).

##### 4.9.4.1 Selectable Markers

Many selectable marker genes may be used in connection with the prostate specific transglutaminase promoter region of the present invention. These include, but are not limited to, a *neo* gene (Potrykus *et al.*, 1985), a nitrilase gene such as *bxn* from *Klebsiella ozaenae* (Stalker *et al.*, 1988), a methotrexate resistant DHFR gene (Thillet *et al.*, 1988), or a mutated

anthranilate synthase gene.

#### 4.9.4.2 Screenable Markers

Screenable markers that may be employed in conjunction with a prostate specific promoter include a  $\beta$ -glucuronidase (GUS) or *uidA* gene, a  $\beta$ -lactamase gene (Sutcliffe, 1978), a *xyIE* gene (Zukowsky *et al.*, 1983), an  $\alpha$ -amylase gene (Ikuta *et al.*, 1990), a tyrosinase gene (Katz *et al.*, 1983), a  $\beta$ -galactosidase gene, a luciferase (*lux*) gene (Ow *et al.*, 1986), an aequorin gene (Prasher *et al.*, 1985), or a gene encoding for green fluorescent protein (Sheen *et al.*, 1995; Haseloff *et al.*, 1997; Reichel *et al.*, 1996; Tian *et al.*, 1997; WO 97/41228).

One screenable marker contemplated for use in the present invention is firefly luciferase, encoded by the *lux* gene. The presence of the *lux* gene in transformed cells may be detected using, for example, X-ray film, scintillation counting, fluorescent spectrophotometry, low-light video cameras, photon counting cameras or multiwell luminometry. It is envisioned that this system may be developed for screening on tissue culture plates, or even for whole tissue or animal screening.

The gene which encodes green fluorescent protein (GFP) is contemplated as a useful reporter gene (Sheen *et al.*, 1995; Haseloff *et al.*, 1997; Reichel *et al.*, 1996; Tian *et al.*, 1997; WO 97/41228). Expression of green fluorescent protein may be visualized in a cell as fluorescence following illumination by particular wavelengths of light. Where use of a screenable marker gene such as *lux* or GFP is desired, benefit may be realized by creating a gene fusion between the screenable marker gene and a selectable marker gene, for example, a GFP-NPTII gene fusion. This could allow, for example, selection of transformed cells followed by screening of transfected or transgenic tissue culture cells or animals.

#### 4.9.5 Multigene Constructs and IRES

In certain embodiments, internal ribosome binding sites (IRES) elements are used to create multigene, or polycistronic, messages. IRES elements are able to bypass the ribosome scanning mode of 5' methylated Cap dependent translation and to begin translation at internal sites (Pelletier and Sonenberg, 1988; Jang *et al.*, 1989). IRES elements from two members of the picornavirus family (polio and encephalomyocarditis) have been described (Pelletier and Sonenberg, 1988), as well an IRES from a mammalian message (Macejak and Sarnow, 1991). IRES elements can be linked to heterologous open reading frames. Multiple open reading frames can be transcribed together, each separated by an IRES, creating polycistronic

messages. By virtue of the IRES element, each open reading frame is accessible to ribosomes for efficient translation. Multiple genes can be efficiently expressed using a single promoter/enhancer to transcribe a single message. Any heterologous open reading frame can be linked to IRES elements. In this way, expression of several proteins can be simultaneously engineered into a cell with a single construct and a single selectable marker.

#### 4.10 Exogenous Genes for Modification of Phenotypes

A particularly important advance of the present invention is that it provides methods and compositions for the efficient expression of selected genes in tissue culture or animal cells. In particular, the current invention provides a prostate specific transglutaminase promoter region for the prostate specific expression of genes in, for example, human subjects.

The choice of a gene for expression in a tissue culture cell line or human host cell in accordance with the invention will depend on the purpose of the transformation. One of the major purposes of transformation is to regulate the growth of or to kill prostate cancer cells. In one embodiment, the expressed gene would be directly toxic to the cell. Such genes would include, for example, the diphtheria toxin A gene (Massuda *et al.*, 1997) or other cytotoxic genes. In another embodiment, the expressed gene may regulate another gene directly or indirectly. The regulated gene may be either endogenously expressed, or expressed under the control of an exogenous promoter introduced by transfection. Such gene may encode toxins, antigens, tumor suppressors, or any other expressed message that would be detrimental to a cell, to slow or kill the cancer. In yet another embodiment, the gene may be therapeutic, and correct a defect that promotes the growth of the cancer cell by, for example, providing a functional copy of a tumor suppressor gene product.

Alternatively, the expressed gene may indirectly kill the prostate cancer through induction of an immune response targeted to the prostate cancer cells. This may be accomplished within the scope of the invention by, for example, providing an expression vector that produces a humanized antibody targeted against a surface-expressed antigen present on the prostate cancer cell. In another embodiment, the expression vector functions as a genetic vaccine to produce an antigenic protein or peptide. By using a prostate specific promoter in the expression vector, only cells of prostate origin should express the antigenic protein or peptide. In a further embodiment, the antigen is selected from surface expressed proteins or peptides known to be up-regulated in prostate cancer cells.

In certain embodiments of the invention, transformation of a recipient cell may be carried out with more than one exogenous (selected) gene. As used herein, an "exogenous gene" or "selected gene" is a gene not normally found in the host genome in an identical context. By this, it is meant that the gene may be isolated from a different species than that of the host genome, or alternatively, isolated from the host genome, but then operably linked to one or more regulatory regions that differ from those found in the unaltered, native gene.

DNA may be introduced into cell lines or individuals for the purpose of expressing RNA transcripts that are not translated into protein. Two examples are antisense RNA and RNA with ribozyme activity. Both may serve possible functions in reducing or eliminating expression of native or introduced genes. However, as detailed below, DNA need not be expressed to affect phenotype.

#### 4.10.1 Antisense RNA

Genes may be constructed or isolated that produce antisense RNA that is complementary to all or part(s) of a targeted messenger RNA(s). The antisense RNA reduces production of the polypeptide product of the messenger RNA. The polypeptide product may be any protein encoded by the genome. The aforementioned genes will be referred to as antisense genes. An antisense gene may be introduced into a tissue culture cell line or individual by transformation methods to reduce expression of a selected protein of interest. For example, antisense constructs to the c-Myc gene may be of therapeutic use for treatment of certain forms of cancer. A number of other oncogenes have been characterized and may serve as potential targets for antisense therapy in cancer. Construction of genes encoding antisense RNAs is well within the skill of the ordinary practitioner, once a suitable target sequence is known.

The term "antisense construct" is intended to refer to nucleic acids, preferably oligonucleotides, that are complementary to the base sequences of a target DNA or RNA. Targeting double-stranded (ds) DNA with an antisense construct leads to triple-helix formation. Targeting RNA will lead to double-helix formation. Antisense nucleic acids, when introduced into a target cell, specifically bind to their target polynucleotide and interfere with transcription, RNA processing, RNA transport, translation and/or stability. Antisense RNA constructs, or DNA encoding such antisense RNAs, may be employed to inhibit gene transcription or translation or both within cells.

Antisense constructs may be designed to bind to complementary sequences within the promoter region or other control regions, exons, introns or even exon-intron boundaries of a



gene. Antisense RNA constructs, or DNA encoding such antisense RNAs, may be employed to inhibit gene transcription or translation or both within a host cell, either *in vitro* or *in vivo*, such as within a human subject. Nucleic acid sequences which comprise "complementary sequences" are those which are capable of base-pairing according to the standard Watson-Crick complementarity rules. Inclusion of less common bases such as inosine, 5-methylcytosine, 6-methyladenine, hypoxanthine and others in hybridizing sequences does not interfere with base-pairing.

As used herein, the term "complementary" means nucleic acid sequences that are substantially complementary over their entire length and have very few base mismatches. For example, nucleic acid sequences of fifteen bases in length may be termed complementary when they have a complementary nucleotide at thirteen or fourteen positions with only a single mismatch. Naturally, nucleic acid sequences which are "completely complementary" will be nucleic acid sequences which are entirely complementary throughout their entire length and have no base mismatches.

Other sequences with lower degrees of homology also are contemplated. For example, an antisense construct that has limited regions of high homology, but also contains a non-homologous region (*e.g.*, a ribozyme) could be designed. These molecules, though having less than 50% homology, would bind to target sequences under appropriate conditions.

While all or part of the gene sequences may be employed in the context of antisense construction, short oligonucleotides are easier to make and increase *in vivo* accessibility. However, both binding affinity and sequence specificity of an antisense oligonucleotide to its complementary target increases with increasing length. One can readily determine whether a given antisense nucleic acid is effective at targeting of the corresponding host cell gene by testing the constructs *in vitro* to determine whether the function of the endogenous gene is affected or whether the expression of related genes having complementary sequences is affected.

In certain embodiments, one may wish to employ antisense constructs that include other elements, for example, those which include C-5 propyne pyrimidines. Oligonucleotides that contain C-5 propyne analogues of uridine and cytidine have been shown to bind RNA with high affinity and to be potent antisense inhibitors of gene expression.

#### 4.10.2 Ribozymes

Another method for inhibiting target gene expression contemplated in the present invention is *via* ribozymes. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). A large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind *via* specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U.S. Patent No. 5,354,855 reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990; Sioud *et al.*, 1992). It has been reported that ribozymes elicit genetic changes in some cells lines to which they were applied. The altered genes included the oncogenes H-ras, c-fos and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Several different ribozyme motifs have been described with RNA cleavage activity (Symons, 1992). Examples that are expected to function equivalently for the down regulation of genes in prostate cancer cells include sequences from the Group I self splicing introns. These include Tobacco Ringspot Virus (Prody *et al.*, 1986), Avocado Sunblotch Viroid (Palukaitis *et al.*, 1979; Symons, 1981), and Lucerne Transient Streak Virus (Forster and Symons, 1987). Sequences from these and related viruses are referred to as hammerhead ribozymes based on a predicted folded secondary structure.

Other suitable ribozymes include sequences from RNase P (Yuan *et al.*, 1992, Yuan and Altman, 1994, U.S. Patent Nos. 5,168,053 and 5,624,824), hairpin ribozyme structures (Berzal-Herranz *et al.*, 1992; Chowrira *et al.*, 1993) and Hepatitis Delta virus based ribozymes (U.S. Patent No. 5,625,047). The general design and optimization of ribozyme directed RNA

cleavage activity has been discussed in detail (Haseloff and Gerlach, 1988, Symons, 1992, Chowrira *et al.*, 1994; Thompson *et al.*, 1995).

The other variable on ribozyme design is the selection of a cleavage site on a given target RNA. Ribozymes are targeted to a given sequence by virtue of annealing to a site by  
5 complimentary base pair interactions. Two stretches of homology are required for this targeting. These stretches of homologous sequences flank the catalytic ribozyme structure defined above. Each stretch of homologous sequence can vary in length from 7 to 15 nucleotides. The only requirement for defining the homologous sequences is that, on the target  
10 RNA, they are separated by a specific sequence that is the cleavage site. For hammerhead ribozymes, the cleavage site is a dinucleotide sequence on the target RNA -- a uracil (U) followed by either an adenine, cytosine or uracil (A,C or U) (Perriman *et al.*, 1992; Thompson *et al.*, 1995). The frequency of this dinucleotide occurring in any given RNA is statistically 3 out of 16.

The large number of possible cleavage sites in, for example, the gene products of  
15 prostate specific transglutaminase, cytokeratin 15, and semenogelin II coupled with the growing number of sequences with demonstrated catalytic RNA cleavage activity indicates that a large number of ribozymes that have the potential to downregulate prostate specific transglutaminase, cytokeratin 15, and semenogelin II are available. Additionally, due to the sequence variation among the prostate specific transglutaminase, cytokeratin 15, and  
20 semenogelin II, ribozymes could be designed to specifically cleave prostate specific transglutaminase, cytokeratin 15, or semenogelin II. Designing and testing ribozymes for efficient cleavage of a target RNA is a process well known to those skilled in the art. Examples of scientific methods for designing and testing ribozymes are described by Chowrira *et al.*, (1994) and Lieber and Strauss (1995), each incorporated by reference. The identification of  
25 operative and preferred sequences for use in prostate specific transglutaminase, cytokeratin 15, and semenogelin II-targeted ribozymes is simply a matter of preparing and testing a given sequence, and is a routinely practiced "screening" method known to those of skill in the art.

The skilled artisan will realize that target sequences for ribozyme degradation are not limited within the scope of the present invention to prostate specific transglutaminase, cytokeratin 15, and semenogelin II. Given the availability of a prostate specific promoter to  
30 selectively express ribozyme sequences, it is considered that virtually any RNA sequence that is essential for cell function could be targeted for ribozyme mediated degradation. It is

contemplated that oncogene messenger RNAs would provide a preferred target for ribozyme degradation in the treatment of cancer.

#### 4.11 Genetic Vaccines

5 This method has been discussed by U.S. Patent No. 5,830,876 to Weiner *et al.*, the entire text of which is incorporated herein by reference. Traditional vaccines have utilized injections of, for example, attenuated or killed viruses to induce an immune response to a pathogenic organism. The possibility of recombination of an attenuated virus into a more virulent form, or incomplete killing of a highly pathogenic organism, poses problems with this  
10 approach. As an alternative, a gene encoding an antigenic, surface-displayed protein or peptide may be genetically engineered into an expression vector and administered to a host animal. By using such a genetic vaccine, the possibility of an inadvertant pathogenic response is eliminated. Such genetic vaccines may be used to induce an immune response against viruses, bacteria or even cancer cells (Weiner and Kennedy, 1999; Ulmer *et al.*, 1993; Barry *et al.*, 1995;  
15 Roman *et al.*, 1997; Boyer and Weiner, 1998; Robinson *et al.*, 1999, the relevant portions of which are incorporated herein by reference).

The present invention allows for the design of genetic vaccines that are preferentially expressed in cells of prostate origin, using the prostate specific promoters disclosed herein. Use of promoters for genes that are over-expressed in prostate cancer cells will allow for an  
20 even more specific targeting of the immune response. In principal, a wide variety of antigenic proteins or peptides could be operably linked to a prostate specific promoter, incorporated into expression vectors and administered to an individual with prostate cancer. Since the specificity of the promoter would limit expression to cells of prostate origin, such as metastatic prostate cancer, only such cells should be attacked by the host immune system. It is contemplated  
25 within the scope of the invention that the gene to be expressed could encode a protein or peptide that naturally occurs on prostate cancer cells, such as carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveollin, POV1, HER2/neu or p27KIP1. In principal, a genetic vaccine targeted against a prostate cancer specific antigen would not only serve to reduce or eliminate existing tumors, but could  
30 also protect the host against future incidence of tumors expressing the same antigenic protein or peptide.

Alternatively, the expressed gene could encode a protein that is not found in the normal host. Since expression of the antigenic protein would be limited to cells of prostate origin, the

resulting immune response would also be limited to such cells. Such encoded proteins could include, for example, hepatitis B surface antigen, herpes simplex glycoprotein, HIV envelope proteins, hemagglutinin, or malaria circumsporozoite protein. Induction of an immune response could also be enhanced by co-expression of cytokine or chemokine genes, operably  
5 linked to a prostate specific promoter (Weiner and Kennedy, 1999).

The vector to be used for construction of the genetic vaccine may comprise any of the expression vectors described herein, such as plasmids, viruses, cosmids, *etc.* The only requirement for the vector to be used is that it must be capable of being transformed into a host cell and must express the encoded protein or peptide in a form and amount that is capable of  
10 provoking an immune response. For certain applications, viral vectors that can infect cells *in situ* in an individual with prostate cancer are preferred. In specific embodiments, a leader sequence may be added to a gene encoding an antigenic protein or peptide that is not normally localized to the cell surface. The selection and design of leader sequences allowing for cell surface targeting is well known in the art. The expression vector may also be designed to  
15 optimize immune reactivity. For example, CG (cytosine:guanine) sequences in the vector may be flanked by two purine residues (adenine or guanine) to their "C" side and two pyrimidines (thymine or cytosine) to their "G" side (Weiner and Kennedy, 1999). Such immunostimulatory sequences provoke a more effective immune response.

#### 20 4.12 Incorporation of DNA into a Target Cell

In certain embodiments of the invention, the nucleic acid encoding the prostate specific transglutaminase promoter region and/or other selected nucleic acid may be stably integrated into the genome of the cell. In yet further embodiments, the nucleic acid(s) may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or  
25 "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed. Preferred vectors are ones designed as "gene therapy vectors" for the treatment of cancer.

30 It is contemplated that the promoter and/or enhancer regions contained within the vectors described herein may be deleted, mutated, or altered to allow the prostate specific transglutaminase promoter region to confer tissue specific or regulated expression of a selected nucleotide sequence. It is also contemplated that the promoter and or enhancer regions of the

vectors described herein may be operatively linked to the prostate specific transglutaminase promoter region to combine regulatory or expression features of a dual or hybrid promoter thus created.

#### 5    4.12.1 DNA Delivery Using Viral Vectors

The ability of certain viruses to infect cells or enter cells *via* receptor-mediated endocytosis, and to integrate into host cell genome and express viral genes stably and efficiently have made them attractive candidates for the transfer of foreign genes into mammalian cells. Preferred gene therapy vectors of the present invention will generally be  
10    viral vectors.

Although some viruses that can accept foreign genetic material are limited in the number of nucleotides they can accommodate and in the range of cells they infect, these viruses have been demonstrated to successfully effect gene expression. However, adenoviruses do not integrate their genetic material into the host genome and therefore do not require host  
15    replication for gene expression, making them ideally suited for rapid, efficient, heterologous gene expression. Techniques for preparing replication-defective infective viruses are well known in the art.

Of course, in using viral delivery systems, one will desire to purify the virion sufficiently to render it essentially free of undesirable contaminants, such as defective  
20    interfering viral particles or endotoxins and other pyrogens such that it will not cause any untoward reactions in the cell, animal or individual receiving the vector construct. A preferred means of purifying the vector involves the use of buoyant density gradients, such as cesium chloride gradient centrifugation.

##### 25    4.12.1.1        Adenoviral Vectors

A particular method for delivery of the expression constructs involves the use of an adenovirus expression vector. Although adenovirus vectors are known to have a low capacity for integration into genomic DNA, this feature is counterbalanced by the high efficiency of gene transfer afforded by these vectors. "Adenovirus expression vector" is meant to include  
30    those constructs containing adenovirus sequences sufficient to support packaging of the construct and to ultimately express a tissue-specific transforming construct that has been cloned therein.

The expression vector comprises a genetically engineered form of adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells  
5 does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity.  
10 Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region  
15 (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter region (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during  
20 the late phase of infection, and all the mRNA's issued from this promoter region possess a 5'-tripartite leader (TPL) sequence which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process.  
25 Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1  
30 proteins (E1A and E1B; Graham *et al.*, 1977). Since the E3 region is dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome

(Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kb of DNA. Combined with the approximately 5.5 kb of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kb, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone.

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain the conditional replication-defective adenovirus vector for use in the present invention. This is because Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to



introduce the transforming construct at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus growth and manipulation is known to those of skill in the art, and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*,  $10^9$  to  $10^{11}$  plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1991; Rich *et al.*, 1993). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993). Recombinant adenovirus and adeno-associated virus (see below) can both infect and transduce non-dividing human primary cells.

#### 4.12.1.2 AAV Vectors

Adeno-associated virus (AAV) is an attractive vector system for use in the present invention as it has a high frequency of integration and it can infect nondividing cells, making it useful for delivery of genes into mammalian cells, in tissue culture (Muzyczka, 1992) or *in vivo*. AAV has a broad host range for infectivity (Tratschin *et al.*, 1984; Laughlin *et al.*, 1986; Lebkowski *et al.*, 1988; McLaughlin *et al.*, 1988). Details concerning the generation and use of rAAV vectors are described in U.S. Patent No. 5,139,941 and U.S. Patent No. 4,797,368, each incorporated herein by reference.

Studies demonstrating the use of AAV in gene delivery include LaFace *et al.* (1988); Zhou *et al.* (1993); Flotte *et al.* (1993); and Walsh *et al.* (1994). Recombinant AAV vectors have been used successfully for *in vitro* and *in vivo* transduction of marker genes (Kaplitt *et al.*, 1994; Lebkowski *et al.*, 1988; Samulski *et al.*, 1989; Yoder *et al.*, 1994; Zhou *et al.*, 1994; 5 Hermonat and Muzyczka, 1984; Tratschin *et al.*, 1985; McLaughlin *et al.*, 1988) and genes involved in human diseases (Flotte *et al.*, 1992; Luo *et al.*, 1994; Ohi *et al.*, 1990; Walsh *et al.*, 1994; Wei *et al.*, 1994). Recently, an AAV vector has been approved for phase I human trials for the treatment of cystic fibrosis.

AAV is a dependent parvovirus in that it requires coinfection with another virus (either 10 adenovirus or a member of the herpes virus family) to undergo a productive infection in cultured cells (Muzyczka, 1992). In the absence of coinfection with helper virus, the wild type AAV genome integrates through its ends into human chromosome 19 where it resides in a latent state as a provirus (Kotin *et al.*, 1990; Samulski *et al.*, 1991). rAAV, however, is not restricted to chromosome 19 for integration unless the AAV Rep protein is also expressed 15 (Shelling and Smith, 1994). When a cell carrying an AAV provirus is superinfected with a helper virus, the AAV genome is "rescued" from the chromosome or from a recombinant plasmid, and a normal productive infection is established (Samulski *et al.*, 1989; McLaughlin *et al.*, 1988; Kotin *et al.*, 1990; Muzyczka, 1992).

Typically, recombinant AAV (rAAV) virus is made by cotransfecting a plasmid 20 containing the gene of interest flanked by the two AAV terminal repeats (McLaughlin *et al.*, 1988; Samulski *et al.*, 1989; each incorporated herein by reference) and an expression plasmid containing the wild type AAV coding sequences without the terminal repeats, for example pIM45 (McCarty *et al.*, 1991; incorporated herein by reference). The cells are also infected or transfected with adenovirus or plasmids carrying the adenovirus genes required for AAV helper 25 function. rAAV virus stocks made in such fashion are contaminated with adenovirus which must be physically separated from the rAAV particles (for example, by cesium chloride density centrifugation). Alternatively, adenovirus vectors containing the AAV coding regions or cell lines containing the AAV coding regions and some or all of the adenovirus helper genes could be used (Yang *et al.*, 1994; Clark *et al.*, 1995). Cell lines carrying the rAAV DNA as an 30 integrated provirus can also be used (Flotte *et al.*, 1995).

#### 4.12.1.3 Retr viral Vect rs

Retroviruses have promise as gene delivery vectors due to their ability to integrate their genes into the host genome, transferring a large amount of foreign genetic material, infecting a broad spectrum of species and cell types and of being packaged in special cell-lines (Miller, 5 1992).

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the 10 retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome. These contain strong 15 promoter region and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding a gene of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, 20 and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media (Nicolas and 25 Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

Concern with the use of defective retrovirus vectors is the potential appearance of wild- 30 type replication-competent virus in the packaging cells. This can result from recombination events in which the intact sequence from the recombinant virus inserts upstream from the gag, pol, env sequence integrated in the host cell genome. However, new packaging cell lines are

now available that should greatly decrease the likelihood of recombination (Markowitz *et al.*, 1988; Hersdorffer *et al.*, 1990).

Gene delivery using second generation retroviral vectors has been reported. Kasahara *et al.* (1994) prepared an engineered variant of the Moloney murine leukemia virus, that normally  
5 infects only mouse cells, and modified an envelope protein so that the virus specifically bound to, and infected, human cells bearing the erythropoietin (EPO) receptor. This was achieved by inserting a portion of the EPO sequence into an envelope protein to create a chimeric protein with a new binding specificity.

#### 10 4.12.1.4 Other Viral Vectors

Other viral vectors may be employed as expression constructs in the present invention. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988), sindbis virus, cytomegalovirus and herpes simplex virus may be employed. They offer several attractive features for various mammalian cells (Friedmann,  
15 1989; Ridgeway, 1988; Baichwal and Sugden, 1986; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large  
20 portions of the genome could be replaced with foreign genetic material. Chang *et al.* (1991) recently introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable  
25 CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

In certain further embodiments, the gene therapy vector will be HSV. A factor that makes HSV an attractive vector is the size and organization of the genome. Because HSV is large, incorporation of multiple genes or expression cassettes is less problematic than in other smaller viral systems. In addition, the availability of different viral control sequences with  
30 varying performance (temporal, strength, *etc.*) makes it possible to control expression to a greater extent than in other systems. It also is an advantage that the virus has relatively few spliced messages, further easing genetic manipulations. HSV also is relatively easy to

manipulate and can be grown to high titers. Thus, delivery is less of a problem, both in terms of volumes needed to attain sufficient MOI and in a lessened need for repeat dosings.

#### 4.12.1.5 Modified Viruses

5 In still further embodiments of the present invention, the nucleic acids to be delivered are housed within an infective virus that has been engineered to express a specific binding ligand. The virus particle will thus bind specifically to the cognate receptors of the target cell and deliver the contents to the cell. A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by  
10 the chemical addition of lactose residues to the viral envelope. This modification can permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

Another approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux  
15 *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

#### 4.12.2 Other Methods of DNA Delivery

20 Several non-viral methods for the transfer of expression constructs into cells also are contemplated by the present invention. In one embodiment of the present invention, the expression construct may consist only of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned which physically or chemically permeabilize the cell membrane. Some of these techniques may be successfully adapted for  
25 *in vivo* or *ex vivo* use, as discussed below.

##### 4.12.2.1 Liposome and Nanocapsule-Mediated Transfection

In a further embodiment of the invention, the expression construct may be entrapped in a liposome. In certain embodiments, the use of liposomes and/or nanoparticles is contemplated  
30 for the introduction of the prostate specific transglutaminase promoter region, operably linked genes, stimulators, inhibitors, or gene therapy vectors, including both wild-type and antisense vectors, into host cells. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid

layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh and Bachhawat, 1991). Also contemplated is an expression construct  
5 complexed with Lipofectamine (Gibco BRL).

Liposome-mediated nucleic acid delivery and expression of foreign DNA *in vitro* has been very successful (Nicolau and Sene, 1982; Fraley and Kaplan, 1979; Nicolau *et al.*, 1987). Wong *et al.* (1980) demonstrated the feasibility of liposome-mediated delivery and expression of foreign DNA in cultured chick embryo, HeLa and hepatoma cells.

10 Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4  $\mu$ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

15 In certain embodiments of the invention, the liposome may be complexed with a hemagglutinating virus (HVJ). This has been shown to facilitate fusion with the cell membrane and promote cell entry of liposome-encapsulated DNA (Kaneda *et al.*, 1989). In other embodiments, the liposome may be complexed or employed in conjunction with nuclear non-histone chromosomal proteins (HMG-1) (Kato *et al.*, 1991). In yet further embodiments, the  
20 liposome may be complexed or employed in conjunction with both HVJ and HMG-1. In other embodiments, the delivery vehicle may comprise a ligand and a liposome. Where a bacterial promoter region is employed in the DNA construct, it also will be desirable to include within the liposome an appropriate bacterial polymerase.

Nanocapsules can generally entrap compounds in a stable and reproducible way. To  
25 avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1  $\mu$ m) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention, and such particles may be easily made.

Liposomes interact with cells *via* four different mechanisms: Endocytosis by  
30 phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous

release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or *vice versa*, without any association of the liposome contents. Varying the liposome formulation can alter which mechanism is operative, although more than one may operate at the same time.

5

#### 4.12.2.2 Electroporation

In certain embodiments of the present invention, the expression construct is introduced into the cell *via* electroporation. Electroporation involves the exposure of a suspension of cells and DNA to a high-voltage electric discharge. Transfection of eukaryotic cells using  
10 electroporation has been quite successful. Mouse pre-B lymphocytes have been transfected with human kappa-immunoglobulin genes (Potter *et al.*, 1984), and rat hepatocytes have been transfected with the chloramphenicol acetyltransferase gene (Tur-Kaspa *et al.*, 1986) in this manner.

#### 15 4.12.2.3 Chemical Transfection

In other embodiments of the present invention, the expression construct is introduced into cells using calcium phosphate precipitation. Human KB cells have been transfected with adenovirus 5 DNA (Graham and Van Der Eb, 1973) using this technique. Also in this manner, mouse L(A9), mouse C127, CHO, CV-1, BHK, NIH3T3 and HeLa cells were transfected with  
20 a neomycin marker gene (Chen and Okayama, 1987), and rat hepatocytes were transfected with a variety of marker genes (Rippe *et al.*, 1990).

In another embodiment, the expression construct is delivered into the cell using DEAE-dextran followed by polyethylene glycol. In this manner, reporter plasmids were introduced into mouse myeloma and erythroleukemia cells (Gopal, 1985).

25

#### 4.12.2.4 Particle Bombardment

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell  
30 membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as

tungsten or gold beads. DNA constructs may be attached to the particles by treatment with calcium and spermidine. The DNA coated particles carry the DNA across the plasma membrane where it is dissociated from the particles in an unknown fashion and incorporated into genomic DNA.

5

#### 4.12.2.5 Direct Microinjection or Sonication Loading

Further embodiments of the present invention include the introduction of the expression construct by direct microinjection or sonication loading. Direct microinjection has been used to introduce nucleic acid constructs into *Xenopus* oocytes (Harland and Weintraub, 1985), and LTK<sup>-</sup> fibroblasts have been transfected with the thymidine kinase gene by sonication loading (Fechheimer *et al.*, 1987).

10

#### 4.12.2.6 Adenoviral Assisted Transfection

In certain embodiments of the present invention, the expression construct is introduced into the cell using adenovirus assisted transfection. Increased transfection efficiencies have been reported in cell systems using adenovirus coupled systems (Kelleher and Vos, 1994; Cotten *et al.*, 1992; Curiel, 1994).

15

#### 4.12.2.7 Receptor Mediated Transfection

Other expression constructs that may be employed to deliver the tissue-specific promoter region and transforming construct to the target cells are receptor-mediated delivery vehicles. These take advantage of the selective uptake of macromolecules by receptor-mediated endocytosis that occurs in selected target cells. In view of the cell type-specific distribution of various receptors, this delivery method adds another degree of specificity to the present invention. Specific delivery in the context of mammalian cells is described by Wu and Wu (1993; incorporated herein by reference).

20

25

Certain receptor-mediated gene targeting vehicles comprise a cell receptor-specific ligand and a DNA-binding agent. Others comprise a cell receptor-specific ligand to which the DNA construct to be delivered has been operatively attached. Several ligands have been used for receptor-mediated gene transfer (Wu and Wu, 1987; Wagner *et al.*, 1990; Perales *et al.*, 1994; Myers, EPO 0273085), which establishes the operability of the technique.

30

In other embodiments, the DNA delivery vehicle component of a cell-specific gene targeting vehicle may comprise a specific binding ligand in combination with a liposome. The



nucleic acids to be delivered are housed within the liposome and the specific binding ligand is functionally incorporated into the liposome membrane. The liposome will thus specifically bind to the receptors of the target cell and deliver the contents to the cell. Such systems have been shown to be functional using systems in which, for example, epidermal growth factor (EGF) is used in the receptor-mediated delivery of a nucleic acid to cells that exhibit upregulation of the EGF receptor.

In still further embodiments, the DNA delivery vehicle component of the targeted delivery vehicles may be a liposome itself, which will preferably comprise one or more lipids or glycoproteins that direct cell-specific binding. For example, Nicolau *et al.* (1987) employed lactosyl-ceramide, a galactose-terminal asialganglioside, incorporated into liposomes and observed an increase in the uptake of the insulin gene by hepatocytes. It is contemplated that the tissue-specific transforming constructs of the present invention can be specifically delivered into the target cells in a similar manner.

#### 4.13 DNA Integration, RNA Expression and Inheritance

Genomic DNA may be isolated from any transformed cell to determine the presence of the exogenous gene through the use of techniques well known to those skilled in the art. Intact sequences will not always be present, due to rearrangement or deletion of sequences in the cell.

The presence of DNA elements introduced through the methods of this invention may be determined by polymerase chain reaction (PCR<sup>TM</sup>). Using this technique, discrete fragments of DNA are amplified and detected by gel electrophoresis. This type of analysis permits one to determine whether a gene is present in a stable transformant, but does not prove integration of the introduced gene into the host cell genome. In addition, it is not possible using PCR<sup>TM</sup> techniques to determine whether transformants have exogenous genes introduced into different sites in the genome, *i.e.*, whether transformants are of independent origin. It is contemplated that using PCR<sup>TM</sup> techniques it would be possible to clone fragments of the host genomic DNA adjacent to an introduced gene.

Positive proof of DNA integration into the host genome and the independent identities of transformants may be determined using Southern hybridization. Using this technique specific DNA sequences that were introduced into the host genome and flanking host DNA sequences can be identified. Hence the Southern hybridization pattern of a given transformant serves as an identifying characteristic of that transformant. In addition it is possible through Southern hybridization to demonstrate the presence of introduced genes in high molecular

weight DNA, *i.e.*, confirm that the introduced gene has been integrated into the host cell genome. The technique of Southern hybridization provides information that is obtained using PCR™, *e.g.*, the presence of a gene, but also demonstrates integration into the genome and characterizes each individual transformant.

5

#### 4.14 Assays of Transgene Expression

Assays may be employed with the instant invention for determination of the relative efficiency of transgene expression. For example, assays may be used to determine the efficacy of deletion mutants of the prostate specific transglutaminase promoter region in directing expression of exogenous genes. Similarly, one could produce random or site-specific mutants of the prostate specific transglutaminase promoter region of the invention and assay the efficacy of the mutants in the expression of a given transgene. Alternatively, assays could be used to determine the function of the prostate specific transglutaminase promoter region in enhancing gene expression when used in conjunction with various different regulatory elements, enhancers, and exogenous genes.

The biological sample to be assayed may comprise nucleic acids isolated from the cells according to standard methodologies (Sambrook *et al.*, 1989). The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to convert the RNA to a complementary DNA. In one embodiment of the invention, the RNA is whole cell RNA; in another, it is poly-A RNA. Normally, the nucleic acid is amplified.

Depending on the format, the specific nucleic acid of interest is identified in the sample directly using. Next, the identified product is detected. In certain applications, the detection may be performed by visual means (*e.g.*, ethidium bromide staining of a gel). Alternatively, the detection may involve indirect identification of the product *via* chemiluminescence, radioactive scintigraphy of radiolabel or fluorescent label or even *via* a system using electrical or thermal impulse signals (Affymax Technology; Bellus, 1994).

Following detection, one may compare the results seen in a given cell line or individual with a statistically significant reference group of non-transformed control cells. In this way, it is possible to detect differences in the amount or kind of mRNA, for example, detected in various transformed tissue culture cell line or animals.

As indicated, a variety of different assays are contemplated in the screening of cells of the current invention. These techniques may in cases be used to detect for both the presence and expression of the particular genes as well as rearrangements that may have occurred in the

gene construct. The techniques include but are not limited to, fluorescent *in situ* hybridization (FISH), direct DNA sequencing, pulsed field gel electrophoresis (PFGE) analysis, Southern or Northern blotting, single-stranded conformation analysis (SSCA), RNase protection assay, allele-specific oligonucleotide (ASO), dot blot analysis, denaturing gradient gel electrophoresis, RFLP and PCR<sup>TM</sup>-SSCP, and isothermal amplification reactions.

#### 4.14.1 Quantitation of Gene Expression with Relative Quantitative RT-PCR<sup>TM</sup>

Reverse transcription (RT) of RNA to cDNA followed by relative quantitative PCR<sup>TM</sup> (RTRQ-PCR<sup>TM</sup>) can be used to determine the relative concentrations of specific mRNA species isolated from tissue culture cell line or animals. By determining that the concentration of a specific mRNA species varies, it is shown that the gene encoding the specific mRNA species is differentially expressed. In this way, a promoters expression profile can be rapidly identified, as can the efficacy with which the promoter directs transgene expression. A more detailed discussion of relative quantitative RT-PCR<sup>TM</sup> is provided in Example 4 below.

#### 4.14.2 Marker Gene Expression

Marker genes represent an efficient means for assaying promoter activity. Using, for example, a selectable marker gene, one could quantitatively determine the resistance conferred upon a tissue culture cell line or animal cell by a construct comprising the selectable marker gene operably linked to the promoter to be assayed, *e.g.* a prostate specific transglutaminase promoter region. Alternatively, various tissue culture cell line or animal parts could be exposed to a selective agent and the relative resistance provided in these parts quantified, thereby providing an estimate of the tissue specific expression of the promoter.

Screenable markers constitute another efficient means for quantifying the expression of a given transgene. Potentially any screenable marker could be expressed and the marker gene product quantified, thereby providing an estimate of the efficiency with which the promoter directs expression of the transgene. Quantification can readily be carried out using either visual means, or, for example, a photon counting device.

A preferred screenable marker gene assay for use with the current invention constitutes the use of the screenable marker gene  $\beta$ -glucuronidase (GUS). Detection of GUS activity can be performed histochemically using 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc) as the substrate for the GUS enzyme, yielding a blue precipitate inside of cells containing GUS activity. This assay has been described in detail (Jefferson, 1987). The blue coloration can

then be visually scored, and estimates of expression efficiency thereby provided. GUS activity also can be determined by immunoblot analysis or a fluorometric GUS specific activity assay (Jefferson, 1987).

Promoter efficiency may also be determined using various techniques directed to quantitation of the amount of protein produced from an operably linked gene. Proteins may be purified and quantitated using methods described above. Alternatively, proteins may be quantified using enzyme activity assays that are well known in the art.

#### 4.15 Therapeutics

One embodiment of the present invention concerns therapy of metastatic prostate cancer by provision of genes encoding therapeutic proteins or peptides operably linked to a prostate specific promoter, such as the prostate specific transglutaminase promoter. In one aspect, the therapeutic gene may encode prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II. Provision of a wild-type gene to an individual with a deficiency of such proteins may be useful in the therapy of prostate cancer.

In alternative aspects, where the levels or activity of prostate specific transglutaminase, cytokeratin 15, and/or semenogelin II is too high, then inhibition of these proteins or the genes encoding them would be adopted as a therapeutic strategy. Inhibitors would include any molecule that reduces the activity or amounts of such proteins or the mRNAs encoding them, including antisense constructs, ribozymes and the like, as well as small molecule inhibitors.

It is contemplated within the scope of the present invention that therapy may be provided by a variety of genes operably linked to a prostate specific transglutaminase promoter or other prostate specific promoter, as described below.

##### 4.15.1 Gene Therapy

The general approach to the aspects of the present invention concerning metastatic prostate cancer therapeutics is to provide a cell with a prostate specific transglutaminase promoter region operably linked to a specific gene, thereby permitting the regulation of the activity of selected proteins or expressed message to take effect. In providing tumor suppressor genes, suicide genes, antisense, ribozymes and other inhibitors under the direction of a prostate specific transglutaminase promoter region, the preferred mode is to provide a nucleic acid encoding the construct to the cell. All such approaches are herein encompassed within the term "gene therapy". Gene therapy vectors incorporating tissue specific promoter elements to express toxins or reporter

genes have been described, and methods of making such constructs would be recognized by one of skill in the art in light of the present disclosures (e.g. see for example, Sambrook *et al.*, 1989; Massuda *et al.*, 1997; Vile and Hart, 1993; Gotoh *et al.*, 1998;

#### 5    4.15.2 Tumor Suppressor Genes

Examples of tumor suppressor genes and candidate tumor suppressor genes contemplated for use in the present invention include, but are not limited to, the retinoblastoma (RB) gene (Friend *et al.*, 1986; Fung *et al.*, 1987; Lee *et al.*, 1987a), the wild-type p53 gene (Finlay *et al.*, 1989; Baker *et al.*, 1990), the deleted in colon carcinoma (DCC) gene (Fearon *et al.*, 1990a, b), the neurofibromatosis type 1 (NF-1) gene (Wallace *et al.*, 1990; Viskochil *et al.*, 1990; Cawthon *et al.*, 1990), the Wilms tumor (WT-1) gene (Call *et al.*, 1990; Gessler *et al.*, 1990; Pritchard-Jones *et al.*, 1990), the von Hippel-Lindau (VHL) disease tumor suppressor gene (Duan *et al.*, 1995), the Maspin (Zou *et al.*, 1994), Brush-1 (Schott *et al.*, 1994) and BRCA 1 genes (Miki *et al.*, 1994; Futreal *et al.*, 1994) for breast cancer, the KAI1 gene (Dong *et al.*, 1995), prostate specific tumor antigen (Corr *et al.*, 1994), and the multiple tumor suppressor (MTS), the bcl-2 gene (Israeli *et al.*, 1997), the c-crbB-2 gene (Israeli *et al.*, 1997), the telomerase (Israeli *et al.*, 1997), DD3 (Bussemakers *et al.*, 1996), p21 (Israeli *et al.*, 1997) or p16 gene (Serrano *et al.*, 1993; Kamb *et al.*, 1994).

##### 20    4.15.2.1      Retinoblastoma

Based upon study of the isolated RB cDNA clone, the predicted RB gene product has 928 amino acids and an expected molecular weight of 106 kDa (Lee *et al.*, 1987b). The natural factor corresponding to the predicted RB gene expression product has been identified as a nuclear phosphoprotein having an apparent relative molecular mass ( $M_r$ ) of between 105 and 114 kDa (Lee *et al.*, 1987b). Various mutations of the RB gene are known, and these are generally inactive. Mutations in RB are seen in virtually all cases of retinoblastoma; additionally, the RB gene products could potentially be inactivated by hyperphosphorylation, and by viral oncoprotein-like cellular protein binding. Although the RB gene was initially named because deletions or mutations within the gene caused the rare childhood ocular tumor, retinoblastoma, loss of pRB function is not only causally related to the retinoblastoma, but is also linked to the progression of many common human cancers. Additionally, there is growing evidence suggesting that the RB protein status is potentially a prognostic marker in urothelial

carcinoma, non-small cell lung carcinoma, and perhaps also in some other types of human neoplasms.

The most direct proof that the cloned RB gene is indeed a tumor suppressor gene comes from introduction of a cloned intact copy of the gene into cancer cells with observed tumor suppression function. A number of reports have indicated that replacement of the normal RB gene in RB-defective tumor cells from disparate types of human cancers could suppress their tumorigenic activity in nude mice (Huang *et al.*, 1988; Goodrich and Lee, 1993; Zhou *et al.*, 1994b). The tumor cell lines studied were derived from widely disparate types of human cancers such as the retinoblastoma, osteosarcoma, carcinomas of the bladder, prostate, breast and lung.

#### 4.15.2.2 p53

Somatic cell mutations of the p53 gene are said to be the most frequent mutations in human cancer (Weinberg, 1991). The normal or wild-type p53 gene is a negative regulator of cell growth, which, when damaged, favors cell transformation (Weinberg, 1991). Tumor cell lines deleted for p53 have been successfully treated with wild-type p53 vector to reduce tumorigenicity (Baker *et al.*, 1990). The p53 gene accumulation has been associated with a subset of prostate cancers (Yang *et al.*, 1996). The inventors contemplate that a wild-type p53 gene, an anti-sense p53 message or ribozyme under the control of a prostate specific transglutaminase promoter region would be a preferred vector construct.

#### 4.15.2.3 DCC

The multiple steps in the tumorigenesis of colon cancer are readily monitored during development by colonoscopy. The combination of colonoscopy with the biopsy of the involved tissue has uncovered a number of degenerative genetic pathways leading to the result of a malignant tumor. One well studied pathway begins with large polyps in which 60% of the cells carry a mutated, activated allele of K-ras. A majority of these tumors then proceed to the inactivation-mutation of the gene referred to as the deleted in colon carcinoma (DCC) gene, followed by the inactivation of the p53 tumor suppressor gene.

The DCC gene is a more than approximately one million base pair gene coding for a 190-kD transmembrane phosphoprotein which is hypothesized to be a receptor (Weinberg, 1991), the loss of which allows the affected cell a growth advantage. It has also been noted that

the DCC has partial sequence homology to the neural cell adhesion molecule (Marshall, 1991) which might suggest a role for the DCC protogene in regulating cell to cell interactions.

#### 4.15.2.4 KAI1

5 Human chromosome 11p11.2-13 has been reported to contain a locus that will suppress metastasis in rat prostatic carcinoma cells. A gene cloned from this locus, KAI1, was shown to suppress metastasis after transfection (Dong *et al.*, 1995). The inventors contemplate this gene will be useful in gene therapy vectors operatively linked to a prostate specific transglutaminase promoter region.

10

#### 4.15.2.5 Prostate specific Membrane Antigen

Prostate specific Membrane Antigen (PSM) also maps to the human chromosome 11p11.2-13 locus reported to suppress metastasis. Full-length PSM cDNA transfected into PC-3 cells produced fewer and smaller tumors when implanted in nude mice (Corr *et al.*, 1994).

15 The inventors contemplate PSMA will be useful in gene therapy vectors operatively linked to a prostate specific transglutaminase promoter region.

#### 4.15.3 Cytokine Genes

20 A number of cytokine genes are contemplated for use in the present invention. Below is an exemplary, but in no way limiting, table of cytokine genes that could be used in certain embodiments of the present invention.

**TABLE 3**

CYTOKINE	REFERENCE
human IL-1 $\alpha$	March <i>et al.</i> , <i>Nature</i> , 315:641, 1985
murine IL-1 $\alpha$	Lomedico <i>et al.</i> , <i>Nature</i> , 312:458, 1984
human IL-1 $\beta$	March <i>et al.</i> , <i>Nature</i> , 315:641, 1985; Auron <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 81:7907, 1984
murine IL-1 $\beta$	Gray, <i>J. Immunol.</i> , 137:3644, 1986; Telford, <i>NAR</i> , 14:9955, 1986
human IL-1ra	Eisenberg <i>et al.</i> , <i>Nature</i> , 343:341, 1990
human IL-2	Taniguchi <i>et al.</i> , <i>Nature</i> , 302:305, 1983; Maeda <i>et al.</i> , <i>Biochem. Biophys. Res. Commun.</i> , 115:1040, 1983

Table 3 – C ntinued

human IL-2	Taniguchi <i>et al.</i> , <i>Nature</i> , 302:305, 1983
human IL-3	Yang <i>et al.</i> , <i>Cell</i> , 47:3, 1986
murine IL-3	Yokota <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 81:1070, 1984; Fung <i>et al.</i> , <i>Nature</i> , 307:233, 1984; Miyatake <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 82:316, 1985
human IL-4	Yokota <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 83:5894, 1986
murine IL-4	Norma <i>et al.</i> , <i>Nature</i> , 319:640, 1986; Lee <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 83:2061, 1986
human IL-5	Azuma <i>et al.</i> , <i>Nuc. Acids Res.</i> , 14:9149, 1986
murine IL-5	Kinashi <i>et al.</i> , <i>Nature</i> , 324:70, 1986; Mizuta <i>et al.</i> , <i>Growth Factors</i> , 1:51, 1988
human IL-6	Hirano <i>et al.</i> , <i>Nature</i> , 324:73, 1986
murine IL-6	Van Snick <i>et al.</i> , <i>Eur. J. Immunol.</i> , 18:193, 1988
human IL-7	Goodwin <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 86:302, 1989
murine IL-7	Namen <i>et al.</i> , <i>Nature</i> , 333:571, 1988
human IL-8	Schmid <i>et al.</i> , <i>J. Immunol.</i> , 139:250, 1987; Matsushima <i>et al.</i> , <i>J. Exp. Med.</i> , 167:1883, 1988; Lindley <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 85:9199, 1988
human IL-9	Renauld <i>et al.</i> , <i>J. Immunol.</i> , 144:4235, 1990
murine IL-9	Renauld <i>et al.</i> , <i>J. Immunol.</i> , 144:4235, 1990
human Angiogenin	Kurachi <i>et al.</i> , <i>Biochemistry</i> , 24:5494, 1985
human GRO $\alpha$	Richmond <i>et al.</i> , <i>EMBO J.</i> , 7:2025, 1988
murine MIP-1 $\alpha$	Davatilis <i>et al.</i> , <i>J. Exp. Med.</i> , 167:1939, 1988
murine MIP-1 $\beta$	Sherry <i>et al.</i> , <i>J. Exp. Med.</i> , 168:2251, 1988
human MIF	Weiser <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 86:7522, 1989
human G-CSF	Nagata <i>et al.</i> , <i>Nature</i> , 319:415, 1986; Souza <i>et al.</i> , <i>Science</i> , 232:61, 1986
human GM-CSF	Cantrell <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 82:6250, 1985; Lee <i>et al.</i> , <i>Proc. Natl. Acad. Sci. USA</i> , 82:4360, 1985; Wong <i>et al.</i> , <i>Science</i> , 228:810, 1985
murine GM-CSF	Gough <i>et al.</i> , <i>EMBO J.</i> , 4:645, 1985



Table 3 – Continued

human M-CSF	Wong, <i>Science</i> , 235:1504, 1987; Kawasaki, <i>Science</i> , 230:291, 1985; Ladner, <i>EMBO J.</i> , 6:2693, 1987
human EGF	Smith <i>et al.</i> , <i>Nuc. Acids Res.</i> , 10:4467, 1982; Bell <i>et al.</i> , <i>NAR</i> , 14:8427, 1986
human TGF- $\alpha$	Derynck <i>et al.</i> , <i>Cell</i> , 38:287, 1984
human FGF acidic	Jaye <i>et al.</i> , <i>Science</i> , 233:541, 1986; Gimenez-Gallego <i>et al.</i> , <i>Biochem. Biophys. Res. Commun.</i> , 138:611, 1986; Harper <i>et al.</i> , <i>Biochem.</i> , 25:4097, 1986
human $\beta$ -ECGF	Jaye <i>et al.</i> , <i>Science</i> , 233:541, 1986
human FGF basic	Abraham <i>et al.</i> , <i>EMBO J.</i> , 5:2523, 1986; Sommer <i>et al.</i> , <i>Biochem. Biophys. Res. Comm.</i> , 144:543, 1987
murine IFN- $\beta$	Higashi <i>et al.</i> , <i>J. Biol. Chem.</i> , 258:9522, 1983; Kuga, <i>NAR</i> , 17:3291, 1989
human IFN- $\gamma$	Gray <i>et al.</i> , <i>Nature</i> , 295:503, 1982; Devos <i>et al.</i> , <i>NAR</i> , 10:2487, 1982; Rinderknecht, <i>J. Biol. Chem.</i> , 259:6790, 1984
human IGF-I	Jansen <i>et al.</i> , <i>Nature</i> , 306:609, 1983; Rotwein <i>et al.</i> , <i>J. Biol. Chem.</i> , 261:4828, 1986
human IGF-II	Bell <i>et al.</i> , <i>Nature</i> , 310:775, 1984
human $\beta$ -NGF chain	Ullrich <i>et al.</i> , <i>Nature</i> , 303:821, 1983
human PDGF A chain	Betsholtz <i>et al.</i> , <i>Nature</i> , 320:695, 1986
human PDGF B chain	Johnsson <i>et al.</i> , <i>EMBO J.</i> , 3:921, 1984; Collins <i>et al.</i> , <i>Nature</i> , 316:748, 1985
human TGF- $\beta$ 1	Derynck <i>et al.</i> , <i>Nature</i> , 316:701, 1985
human TNF- $\alpha$	Pennica <i>et al.</i> , <i>Nature</i> , 312:724, 1984; Fransen <i>et al.</i> , <i>Nuc. Acids Res.</i> , 13:4417, 1985
human TNF- $\beta$	Gray <i>et al.</i> , <i>Nature</i> , 312:721, 1984
murine TNF- $\beta$	Gray <i>et al.</i> , <i>Nucl. Acids Res.</i> , 15:3937, 1987

#### 4.15.4 Cytotoxic Genes

5 Genes that are toxic to cells can be expressed from a prostate specific transglutaminase promoter region, thereby permitting the regulation of the activity of a toxin. The prostate tissue and prostate tissue derived expression of such genes would kill the cells transfected with such a construct, including cancer cells. These type of genes are also known as "suicide genes", and are contemplated as being particularly useful in the treatment of prostate derived cancer. Genes  
10 contemplated for use include, but are not limited to, the diphtheria toxin A chain (Massude *et al.*,

1997).

Diphtheria toxin A chain (DT-A) is a bacterial protein that catalyzes the ADP ribosylation of the diphthamide group of cellular elongation factor 2. This inhibits protein translation and activates apoptosis. It has been shown that expression of this gene under the control of tissue specific promoters in transgenic mice produces tissue specific cell toxicity. An alveolar rhabdomyosarcoma (ARMS) cancer cell line transfected with DT-A expressed from a promoter specifically induced by transcription factors produced by the ARMS was selectively killed by the expressed toxin (Massuda *et al.*, 1997). The inventors contemplate that DT-A operatively linked to a prostate specific transglutaminase promoter region will be useful in killing prostate derived cancers.

#### 4.15.5 Heterologous Genes

While tumor suppressor and cytokine genes are preferred in a number of embodiments of the present invention, in other embodiments a variety of heterologous genes are contemplated for use. Such genes may be operably linked to a prostate specific transglutaminase promoter or other promoter to provide a therapeutic effect in various disease states. Below is a list of selected cloned structural genes that could be used in the present invention. The list is not in any way meant to be interpreted as limiting, only as exemplary of the types of structural genes contemplated for use in the present invention.

**TABLE 4**

SELECTED CLONED STRUCTURAL GENES		
GENE	CLONE TYPE*	REFERENCE
activin	porcine-cDNA	Mason AJ, Nat, 318:659, 1985
adenosine deaminase	h-cDNA	Wiginton DA, PNAS, 80:7481, 1983
angiotensinogen I	r-cDNA	Ohkubo H, PNAS, 80:2196, 1983
	r-gDNA	Tanaka T, JBC, 259:8063, 1984
antithrombin III	H-cDNA	Bock SC, NAR 10:8113, 1982
	h-cDNA and gDNA	Prochownik EV, JBC, 258:8389, 1983

Table 4 – Continued

antitrypsin, alpha I	h-cDNA h-gDNA RFLP	Kurachi K, PNAS, 78:6826, 1981 Leicht M, Nat, 297:655, 1982 Cox DW, AJHG, 36:134S, 1984
apolipoprotein A-I	h-cDNA, h-gDNA RFLP h-gDNA	Shoulders CC, NAR, 10:4873, 1982 Karathanasis SK, Nat, 301:718, 1983 Kranthanas SK, PNAS, 80:6147, 1983
apolipoprotein A-II	h-cDNA Chr h-cDNA	Sharpe CR, NAR, 12:3917, 1984 Sakaguchi, AY, AJHB, 36:207S, 1984 Knott TJ, BBRC, 120:734, 1984
apolipoprotein C-I	h-cDNA	Knott TJ, NAR, 12:3909, 1984
apolipoprotein C-II	h-cDNA h-cDNA h-cDNA RFLP	Jackson CL, PNAS, 81:2945, 1984 Mykelbost O, JBC, 249:4401, 1984 Fojo SS, PNAS, 81:6354, 1984 Humphries SE, C Gen, 26:389, 1984
apolipoprotein C-III	h-cDNA and gDNA h-cDNA	Karathanasis SK, Nat, 304:371, 1983 Sharpe CR, NAR, 12:3917, 1984
apolipoprotein E	h-cDNA	Brewslow JL, JBC, 257:14639, 1982
atrial natriuretic factor	h-cDNA h-cDNA h-cDNA h-gDNA h-gDNA h-gDNA	Oikawa S, Nat, 309:724, 1984 Nakayama K, Nat, 310:699, 1984 Zivin RA, PNAS, 81:6325, 1984 Seidman CE, Sci, 226:1206, 1984 Nemer M, Nat, 312:654, 1984 Greenberg BI, Nat, 312:665, 1984
chorionic gonadotropin,	h-cDNA	Fiddes JC, Nat, 281:351, 1981
alpha chain	RFLP	Boethby M, JBC, 256:5121, 1981
chorionic gonadotropin,	h-cDNA	Fiddes JC, Nat, 286:684, 1980
beta chain	h-gDNA h-gDNA	Boorstein WR, Nat, 300:419, 1982 Talmadge K, Nat, 307:37, 1984
chymosin, pro (rennin)	bovine-cDNA	Harris TJR, NAR, 10:2177, 1982
complement, factor B	h-cDNA h-cDNA and gDNA	Woods DE, PNAS, 79:5661, 1982 Duncan R, PNAS, 80:4464, 1983
complement C2	h-cDNA h-gDNA (C2, C4, and B)	Bentley DR, PNAS, 81:1212, 1984 Carroll MC, Nat, 307:237, 1984
complement C3	m-cDNA h-gDNA	Domdey H, PNAS, 79:7619, 1983 Whitehead AS, PNAS, 79:5021, 1982

Table 4 - Continued

complement C4	h-cDNA and gDNA	Carroll MC, PNAS, 80:264, 1983
	h-cDNA	Whitehead AS, PNAS, 80:5387, 1983
complement C9	h-cDNA	DiScipio RC, PNAS, 81:7298, 1984
corticotropin releasing factor	sheep-cDNA	Furutani Y, Nat, 301:537, 1983
	h-gDNA	Shibahara S, EMBO J, 2:775, 1983
epidermal growth factor	m-cDNA	Gray A, Nat, 303:722, 1983
	m-cDNA	Scott J, Sci, 21:236, 1983
	h-gDNA	Brissenden JE, Nat, 310:781, 1984
epidermal growth factor	h-cDNA and Chr	Lan CR, Sci, 224:843, 1984
receptor, oncogene		
c-erb B		
epoxide dehydratase	r-cDNA	Gonzalez FJ, JBC, 256:4697, 1981
erythropoietin	h-cDNA	Lec-Huang S, PNAS, 81:2708, 1984
esterase inhibitor, C1	h-cDNA	Stanley KK, EMBO J, 3:1429, 1984
factor VIII	h-cDNA and gDNA	Gitschier J, Nat, 312:326, 1984
	h-cDNA	Toole JJ, Nat, 312:342, 1984
factor IX, Christmas factor	h-cDNA	Kutachi K, PNAS, 79:6461, 1982
	h-cDNA	Choo KH, Nat, 299:178, 1982
	RFLP	Camcrino G, PNAS, 81:498, 1984
	h-gDNA	Anson DS, EMBO J, 3:1053, 1984
factor X	h-cDNA	Leytus SP, PNAS, 81:3699, 1984
fibrinogen A alpha,	h-cDNA	Kant JA, PNAS, 80:3953, 1983
B beta, gamma	h-gDNA (gamma)	Fornace AJ, Sci, 224:161, 1984
	h-cDNA (alpha gamma)	Imam AMA, NAR, 11:7427, 1983
	h-gDNA (gamma)	Fornace AJ, JBC, 259:12826, 1984
gastrin releasing peptide	h-cDNA	Spindel ER, PNAS, 81:5699, 1984
glucagon, prepro	hamster c-DNA	Bell GI, Nat, 302:716, 1983
	h-gDNA	Bell GI, Nat, 304:368, 1983
growth hormone	h-cDNA	Martial JA, Sci, 205:602, 1979
	h-gDNA	DeNoto FM, NAR, 9:3719, 1981
	GH-like gene	Owerbach, D, Sci, 209:289, 1980
growth hormone, RF,	h-cDNA	Gubler V, PNAS, 80:3411, 1983
somatocrinin	h-cDNA	Mayo KE, Nat, 306:86:1983
hemopexin	h-cDNA	Stanley KK, EMBO J, 3:1429, 1984
inhibin	porcine-cDNA	Mason AJ, Nat, 318:659, 1985

Table 4 - Continued

insulin, prepro	h-gDNA	Ullrich a, Sci, 209:612, 1980
insulin-like growth factor I	h-cDNA	Jansen M, Nat, 306:609, 1983
	h-cDNA	Bell GI, Nat, 310:775, 1984
	Chr	Brissenden JE, Nat, 310:781, 1984
insulin-like growth factor II	h-cDNA	Bell GI, Nat, 310:775, 1984
	h-gDNA	Dull TJ, Nat, 310:777, 1984
	Chr	Brissenden JE, Nat, 310:781, 1984
interferon, alpha	h-cDNA	Maeda S, PNAS, 77:7010, 1980
(leukocyte), multiple	h-cDNA (8 distinct)	Goeddel DV, Nat., 290:20, 1981
	h-gDNA	Lawn RM, PNAS, 78:5435, 1981
	h-gDNA	Todokoro K, EMBO J, 3:1809, 1984
	h-gDNA	Torczynski RM, PNAS, 81:6451, 1984
interferon, beta (fibroblast)	h-cDNA	Taniguchi T, Gene, 10:11, 1980
	h-gDNA	Lawn RM, NAR, 9:1045, 1981
	h-gDNA (related)	Sehgal P, PNAS, 80:3632, 1983
	h-gDNA (related)	Sagar AD, Sci, 223:1312, 1984
interferon, gamma	h-cDNA	Gray PW, Nat, 295:503, 1982
(immune)	h-gDNA	Gray PW, Nat, 298:859, 1982
interleukin-1	m-cDNA	Lomedico PT, Nat, 312:458, 1984
interleukin-2, T-cell	h-cDNA	Devos R, NAR, 11:4307, 1983
growth factor	h-cDNA	Taniguchi T, Nat, 302:305, 1983
	h-gDNA	Hollbrook NJ, PNAS, 81:1634, 1984
	Chr	Siegel LF, Sci, 223:175, 1984
interleukin-3	m-cDNA	Fung MC, Nat, 307:233, 1984
kininogen, two forms	bovine-cDNA	Nawa H, PNAS, 80:90, 1983
	bovine,-cDNA and gDNA	Kitamura N, Nat, 305:545, 1983
leuteinizing hormone, beta subunit	h-gDNA and Chr	Talmadge K, Nat, 207:37, 1984
leuteinizing hormone	h-cDNA and gDNA	Seeburg PH, Nat, 311:666, 1984
releasing hormone		
lymphotoxin	h-cDNA and gDNA	Gray PW, Nat, 312:721, 1984
mast cell growth factor	m-cDNA	Yokoya T, PNAS, 81:1070, 1984
nerve growth factor, beta subunit	m-cDNA	Scott J, Nat, 302:538, 1983
	h-gDNA	Ullrich A, Nat, 303:821, 1983
	Chr	Franke C, Sci, 222:1248, 1983

Table 4 - C ntinued

oncogene, c-sis, PGDF	h-gDNA	Dalla-Favera R, Nat, 295:31, 1981
chain A	h-cDNA	Clarke MF, Nat, 208:464, 1984
pancreatic polypeptide and icosapeptide	h-cDNA	Boel E, EMBO J, 3:909, 1984
parathyroid hormone,	h-cDNA	Hendy GN, PNAS, 78:7365, 1981
prepro	h-gDNA	Vasicek TJ, PNAS, 80:2127, 1983
plasminogen	h-cDNA and gDNA	Malinowski DP, Fed P, 42:1761, 1983
plasminogen activator	h-cDNA	Edlund T, PNAS, 80:349, 1983
	h-cDNA	Pennica D, Nat, 301:214, 1983
	h-gDNA	Ny T, PNAS, 81:5355, 1984
prolactin	h-cDNA	Cook NE, JBC, 256:4007, 1981
	r-gDNA	Cooke NE, Nat, 297:603, 1982
proopiomelanocortin	h-cDNA	DeBold CR, Sci, 220:721, 1983
	h-gDNA	Cochet M, Nat, 297:335, 1982
protein C	h-cDNA	Foster D, PNAS, 81:4766, 1984
prothrombin	bovine-cDNA	MacGillivray RTA, PNAS, 77:5153, 1980
relaxin	h-gDNA	Hudson P, Nat, 301:628, 1983
	h-cDNA (2 genes)	Hudson P, EMBO J, 3:2333, 1984
	Chr	Crawford, RJ, EMBO J, 3:2341, 1984
renin, prepro	h-cDNA	Imai T, PNAS, 80:7405, 1983
	h-gDNA	Hobart PM, PNAS 81:5026, 1984
	h-gDNA	Miyazaki H, PNAS, 81:5999, 1984
	Chr	Chirgwin JM, SCMG, 10:415, 1984
somatostatin	h-cDNA	Shen IP, PNAS, 79:4575, 1982
	h-gDNA and Ri-IP	Naylot SI, PNAS, 80:2686, 1983
tachykinin, prepro,	bovine-cDNA	Nawa H, Nat, 306:32, 1983
substances P and K	bovine-gDNA	Nawa H, Nat, 312:729, 1984
urokinase	h-cDNA	Verde P, PNAS, 81:4727, 1984
vasoactive intestinal peptide, prepro	h-cDNA	Itoh N, Nat, 304:547, 1983
vasopressin	r-cDNA	Schmale H, EMBO J, 2:763, 1983

Key to Table 4: \*cDNA - complementary DNA; Chr - chromosome; gDNA - genomic DNA;

5 RFLP - restriction fragment polymorphism; h - human; m - mouse; r - rat

#### 4.16 Homologous Recombination

Although genetic transformation tends to be quite efficient, it is also accompanied by problems associated with random insertion. Random integration can lead to the inactivation of essential genes, or to the aberrant expression of the introduced gene. Additional problems associated with genetic transformation include mosaicism due to multiple integrations, and technical difficulties associated with generation of replication defective recombinant viral vectors.

Some of these drawbacks can be overcome by the utilization of a technique known as homologous recombination (Koller and Smithies, 1992). This technique allows the precise modification of existing genes, overcomes the problems of positional effects and insertional inactivation, and allows the inactivation of specific genes, as well as the replacement of one gene for another. Methods for homologous recombination are described in U. S. Patent 5,614,396 and U.S. Patent No. 5,527,695, incorporated herein in their entirety by reference.

Thus a preferred method for the delivery of expression constructs involves the use of homologous recombination. Homologous recombination relies, like antisense, on the tendency of nucleic acids to base pair with complementary sequences. In this instance, the base pairing serves to facilitate the interaction of two separate nucleic acid molecules so that strand breakage and repair can take place. In other words, the "homologous" aspect of the method relies on sequence homology to bring two complementary sequences into close proximity, while the "recombination" aspect provides for one complementary sequence to replace the other by virtue of the breaking of certain bonds and the formation of others.

Put into practice, homologous recombination is used as follows. First, a site for integration is selected within the host cell. Sequences homologous to the integration site are then included in a genetic construct, flanking the selected gene to be integrated into the genome. Flanking, in this context, simply means that target homologous sequences are located both upstream (5') and downstream (3') of the selected gene. These sequences should correspond to some sequences upstream and downstream of the target gene. The construct is then introduced into the cell, thus permitting recombination between the cellular sequences and the construct.

As a practical matter, the genetic construct will normally act as far more than a vehicle to insert the gene into the genome. For example, it is important to be able to select for recombinants and, therefore, it is common to include within the construct a selectable marker gene. This gene permits selection of cells that have integrated the construct into their genomic

DNA by conferring resistance to various biostatic and biocidal drugs. In addition, this technique may be used to "knock-out" (delete) or interrupt a particular gene. Thus, another approach for inhibiting prostate gene expression involves the use of homologous recombination, or "knock-out technology". This is accomplished by including a mutated or  
5 vastly deleted form of the heterologous gene between the flanking regions within the construct.

DNA can be inserted into the host genome by a homologous recombination reaction involving either a single reciprocal recombination (resulting in the insertion of the entire length of the introduced DNA) or through a double reciprocal recombination (resulting in the insertion of only the DNA located between the two recombination events). For example, if one wishes  
10 to insert a foreign gene into the genomic site where a selected gene is located, the introduced DNA should contain sequences homologous to the selected gene. A single homologous recombination event would then result in the entire introduced DNA sequence being inserted into the selected gene. Alternatively, a double recombination event can be achieved by flanking each end of the DNA sequence of interest (the sequence intended to be inserted into  
15 the genome) with DNA sequences homologous to the selected gene. A homologous recombination event involving each of the homologous flanking regions will result in the insertion of the foreign DNA. Thus only those DNA sequences located between the two regions sharing genomic homology become integrated into the genome.

Although introduced sequences can be targeted for insertion into a specific genomic site  
20 *via* homologous recombination, in higher eukaryotes homologous recombination is a relatively rare event compared to random insertion events. In tissue culture cell line or animal cells, foreign DNA molecules find homologous sequences in the cell's genome and recombine at a frequency of approximately  $0.5-4.2 \times 10^{-4}$ . Thus any transformed cell that contains an introduced DNA sequence integrated *via* homologous recombination will also likely contain  
25 numerous copies of randomly integrated introduced DNA sequences. Therefore, to maintain control over the copy number and the location of the inserted DNA, these randomly inserted DNA sequences can be removed. One manner of removing these random insertions is to utilize a site-specific recombinase system. In general, a site specific recombinase system consists of three elements: two pairs of DNA sequence (the site - specific recombination sequences) and a  
30 specific enzyme (the site-specific recombinase). The site-specific recombinase will catalyze a recombination reaction only between two site -specific recombination sequences.

A number of different site specific recombinase systems could be employed in accordance with the instant invention, including, but not limited to, the Cre/lox system of



bacteriophage P1 (U.S. Patent No. 5,658,772, specifically incorporated herein by reference in its entirety), the Gin recombinase of phage Mu (Maeser *et al.*, 1991), the Pin recombinase of *E. coli* (Enomoto *et al.*, 1983), and the R/RS system of the pSR1 plasmid (Araki *et al.*, 1992).

Another refinement of the homologous recombination approach involves the use of a  
5 "negative" selectable marker. One example of the use of the cytosine deaminase gene in a negative selection method is described in U.S. Patent No. 5,624,830. The negative selection marker, unlike the selectable marker, causes death of cells which express the marker. Thus, it is used to identify undesirable recombination events. When seeking to select homologous recombinants using a selectable marker, it is difficult in the initial screening step to identify  
10 proper homologous recombinants from recombinants generated from random, non-sequence specific events. These recombinants also may contain the selectable marker gene and may express the heterologous protein of interest, but will, in all likelihood, not have the desired phenotype. By attaching a negative selectable marker to the construct, but outside of the flanking regions, one can select against many random recombination events that will  
15 incorporate the negative selectable marker. Homologous recombination should not introduce the negative selectable marker, as it is outside of the flanking sequences.

#### 4.17 Pharmaceutical Compositions

##### 20 4.17.1 Pharmaceutically Acceptable Carriers

Another embodiment of the present invention concerns methods for the treatment of cancer. The present invention contemplates the use of compounds having activity to modulate expression from the prostate specific transglutaminase or other promoter regions. Treatment methods will involve treating an individual with an effective amount of such a regulatory  
25 compound. Alternatively, therapy may comprise the administration of a therapeutically effective amount of an expression vector comprising a prostate specific transglutaminase promoter, operably linked to a therapeutic gene.

Aqueous compositions of the present invention comprise an effective amount of the composition of interest, dissolved or dispersed in a pharmaceutically acceptable carrier or  
30 aqueous medium. Aqueous compositions of gene therapy vectors expressing any of the foregoing are also contemplated. The phrases "pharmaceutically or pharmacologically acceptable" refer to molecular entities and compositions that do not produce an adverse, allergic or other untoward reaction when administered to an animal, or a human, as appropriate.

Administration of the compound to a patient will follow general protocols for the administration of chemotherapeutics, taking into account the toxicity, if any, of the compound. It is expected that the treatment cycles would be repeated as necessary.

Aqueous compositions of the present invention comprise an effective amount of the compound, dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium. Such compositions can also be referred to as inocula. As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. For human administration, preparations should meet sterility, pyrogenicity, general safety and purity standards as required by FDA Office of Biologics standards.

The biological material should be extensively dialyzed to remove undesired small molecular weight molecules and/or lyophilized for more ready formulation into a desired vehicle, where appropriate. The active compounds will then generally be formulated for parenteral administration, *e.g.*, formulated for injection *via* the intravenous, intramuscular, subcutaneous, intralesional, or even intraperitoneal routes. The preparation of an aqueous composition that contains a desired agent as an active component or ingredient will be known to those of skill in the art in light of the present disclosure. Typically, such compositions can be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for use in preparing solutions or suspensions upon the addition of a liquid prior to injection can also be prepared; and the preparations can also be emulsified.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions; formulations including sesame oil, peanut oil or aqueous propylene glycol; and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases the form must be sterile and must be fluid. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi.

Solutions of the active compounds as free base or pharmacologically acceptable salts can be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof

and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

A therapeutic agent can be formulated into a composition in a neutral or salt form. Pharmaceutically acceptable salts, include the acid addition salts (formed with the free amino  
5 groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. In terms of  
10 using peptide therapeutics as active ingredients, the technology of U.S. Patents 4,608,251; 4,601,903; 4,599,231; 4,599,230; 4,596,792; and 4,578,770, each incorporated herein by reference, may be used.

The carrier can also be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and  
15 the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin. by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many  
20 cases, it will be preferable to include isotonic agents. for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated  
25 above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder  
30 of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof. The preparation of more, or highly, concentrated solutions for direct injection is also contemplated, where the use of DMSO as solvent is envisioned to result in extremely rapid penetration, delivering high concentrations of the active agents to a small tumor area.

Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms, such as the type of injectable solutions described above, but drug release capsules and the like can also be employed.

5 For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, sterile aqueous media that can be employed will be known to those of skill in the art in light of the  
10 present disclosure. For example, one dosage could be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580).

The term "unit dose" refers to physically discrete units suitable for use in a subject, each  
15 unit containing a predetermined quantity of the therapeutic composition calculated to produce the desired responses, discussed above, in association with its administration, *i.e.*, the appropriate route and treatment regimen. The quantity to be administered, both according to number of treatments and unit dose, depends on the subject to be treated, the state of the subject and the protection desired. The person responsible for administration will, in any event, determine the  
20 appropriate dose for the individual subject.

The active agents may be formulated within a therapeutic mixture to comprise about 0.0001 to 1.0 milligrams, or about 0.001 to 0.1 milligrams, or about 0.1 to 1.0 or even about 10 milligrams per dose or so. Multiple doses can also be administered.

In addition to the compounds formulated for parenteral administration, such as  
25 intravenous or intramuscular injection, other pharmaceutically acceptable forms include, *e.g.*, tablets or other solids for oral administration; liposomal formulations; time release capsules; and any other form currently used.

One may also use nasal solutions or sprays, aerosols or inhalants in the present invention. Nasal solutions are usually aqueous solutions designed to be administered to the nasal passages in  
30 drops or sprays. Nasal solutions are prepared so that they are similar in many respects to nasal secretions, so that normal ciliary action is maintained. Thus, the aqueous nasal solutions usually are isotonic and slightly buffered to maintain a pH of 5.5 to 6.5. In addition, antimicrobial preservatives, similar to those used in ophthalmic preparations, and appropriate drug stabilizers, if

required, may be included in the formulation. Various commercial nasal preparations are known and include, for example, antibiotics and antihistamines and are used for asthma prophylaxis.

Additional formulations suitable for other modes of administration include suppositories and pessaries. A rectal pessary or suppository may also be used. Suppositories are solid dosage  
5 forms of various weights and shapes, usually medicated, for insertion into the rectum or the urethra. After insertion, suppositories soften, melt or dissolve in the cavity fluids. In general, for suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1%-2%.

10 Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders. In certain defined embodiments, oral pharmaceutical compositions will comprise an inert diluent or  
15 assimilable edible carrier, or they may be enclosed in hard or soft shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compounds may be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 0.1%  
20 of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 75% of the weight of the unit, or preferably between 25-60%. The amount of active compounds in such therapeutically useful compositions is such that a suitable dosage will be obtained.

The tablets, troches, pills, capsules and the like may also contain the following: a  
25 binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of  
30 the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compounds

sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor.

#### 4.17.2 Combination Therapies

5 Therapies according to the present invention encompass combination therapies that include treatment with compositions comprising therapeutic expression vectors or small molecule inhibitors or activators of prostate specific promoters, as well as standard chemo- and radiotherapies. For example, chemotherapeutics include, but are not limited to, cisplatin (CDDP), carboplatin, procarbazine, mechlorethamine, cyclophosphamide, ifosfamide, melphalan, 10 chlorambucil, bisulfan, nitrosurea, dactinomycin, daunorubicin, doxorubicin, bleomycin, plicomycin, mitomycin, etoposide (VP16), tamoxifen, taxol, transplatinum, 5-fluorouracil, vincristin, vinblastin and methotrexate. Also included in combined therapies may be x- and  $\gamma$ -irradiation. So long as a particular therapeutic approach is not known to be detrimental in itself, or counteracts the effectiveness of other administered therapies, its combination with the 15 present invention is contemplated.

#### 4.17.3 Radiation Therapy

In terms of surgery, any surgical intervention may be practiced in combination with the present invention. In connection with radiotherapy, any mechanism for inducing DNA damage 20 locally within tumor cells is contemplated, such as  $\gamma$ -irradiation, X-rays, UV-irradiation, microwaves and even electronic emissions and the like. The directed delivery of radioisotopes to tumor cells is also contemplated, and this may be used in connection with a targeting antibody or other targeting means.

#### 25 4.17.4 Chemotherapeutic Agents

Cytokine therapy has proven to be an effective partner for combined therapeutic regimens. Various cytokines may be employed in such combined approaches. Examples of cytokines include IL-1 $\alpha$ , IL-1 $\beta$ , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, TGF- $\beta$ , GM-CSF, M-CSF, G-CSF, TNF $\alpha$ , TNF $\beta$ , LAF, TCGF, BCGF, TRF, 30 BAF, BDG, MP, LIF, OSM, TMF, PDGF, IFN- $\alpha$ , IFN- $\beta$ , IFN- $\gamma$ . Cytokines are administered according to standard regimens, consistent with clinical indications such as the condition of the patient and relative toxicity of the cytokine.

Compositions of the present invention can have an effective amount of an engineered virus or cell for therapeutic administration in combination with an effective amount of a compound (second agent) that is a chemotherapeutic agent as exemplified below. Such compositions will generally be dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium.

A wide variety of chemotherapeutic agents may be used in combination with the therapeutic genes of the present invention. These can be, for example, agents that directly cross-link DNA, agents that intercalate into DNA, and agents that lead to chromosomal and mitotic aberrations by affecting nucleic acid synthesis.

Agents that directly cross-link nucleic acids, specifically DNA, are envisaged and are shown herein, to eventuate DNA damage leading to a synergistic antineoplastic combination. Agents such as cisplatin, and other DNA alkylating agents may be used.

Agents that damage DNA also include compounds that interfere with DNA replication, mitosis, and chromosomal segregation. Examples of these compounds include adriamycin (also known as doxorubicin), VP-16 (also known as etoposide), verapamil, podophyllotoxin, and the like. Widely used in clinical setting for the treatment of neoplasms these compounds are administered through bolus injections intravenously at doses ranging from 25-75 mg/m<sup>2</sup> at 21 day intervals for adriamycin, to 35-100 mg/m<sup>2</sup> for etoposide intravenously or orally.

#### 4.17.4.1 Doxorubicin

Doxorubicin hydrochloride, 5,12-Naphthacenedione, (8*s-cis*)-10-[(3-amino-2,3,6-trideoxy- $\alpha$ -L-lyxo-hexopyranosyl)oxy]-7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-hydrochloride (hydroxydaunorubicin hydrochloride, Adriamycin) is used in a wide antineoplastic spectrum. It binds to DNA and inhibits nucleic acid synthesis, inhibits mitosis and promotes chromosomal aberrations.

Administered alone, it is the drug of first choice for the treatment of thyroid adenoma and primary hepatocellular carcinoma. It is a component of 31 first-choice combinations for the treatment of ovarian, endometrial and breast tumors, bronchogenic oat-cell carcinoma, non-small cell lung carcinoma, gastric adenocarcinoma, retinoblastoma, neuroblastoma, mycosis fungoides, pancreatic carcinoma, prostatic carcinoma, bladder carcinoma, myeloma, diffuse histiocytic lymphoma, Wilms' tumor, Hodgkin's disease, adrenal tumors, osteogenic sarcoma soft tissue sarcoma, Ewing's sarcoma, rhabdomyosarcoma and acute lymphocytic leukemia. It

is an alternative drug for the treatment of islet cell, cervical, testicular and adrenocortical cancers. It is also an immunosuppressant.

Doxorubicin is absorbed poorly and must be administered intravenously. The pharmacokinetics are multicompartmental. Distribution phases have half-lives of 12 min and 3.3 h. The elimination half-life is about 30 h. Forty to 50% is secreted into the bile. Most of the remainder is metabolized in the liver, partly to an active metabolite (doxorubicinol), but a few percent is excreted into the urine. In the presence of liver impairment, the dose should be reduced.

Appropriate doses are, intravenous, adult, 60 to 75 mg/m<sup>2</sup> at 21-day intervals or 25 to 30 mg/m<sup>2</sup> on each of 2 or 3 successive days repeated at 3- or 4-wk intervals or 20 mg/m<sup>2</sup> once a wk. The lowest dose should be used in elderly patients, when there is prior bone-marrow depression caused by prior chemotherapy or neoplastic marrow invasion, or when the drug is combined with other myelopoietic suppressant drugs. The dose should be reduced by 50% if the serum bilirubin lies between 1.2 and 3 mg/dl and by 75% if above 3 mg/dl. The lifetime total dose should not exceed 550 mg/m<sup>2</sup> in patients with normal heart function and 400 mg/m<sup>2</sup> in persons having received mediastinal irradiation. Alternatively, 30 mg/m<sup>2</sup> on each of 3 consecutive days, repeated every 4 wk. Exemplary doses may be 10 mg/m<sup>2</sup>, 20 mg/m<sup>2</sup>, 30 mg/m<sup>2</sup>, 50 mg/m<sup>2</sup>, 100 mg/m<sup>2</sup>, 150 mg/m<sup>2</sup>, 175 mg/m<sup>2</sup>, 200 mg/m<sup>2</sup>, 225 mg/m<sup>2</sup>, 250 mg/m<sup>2</sup>, 275 mg/m<sup>2</sup>, 300 mg/m<sup>2</sup>, 350 mg/m<sup>2</sup>, 400 mg/m<sup>2</sup>, 425 mg/m<sup>2</sup>, 450 mg/m<sup>2</sup>, 475 mg/m<sup>2</sup>, 500 mg/m<sup>2</sup>. Of course, all of these dosages are exemplary, and any dosage in-between these points is also expected to be of use in the invention.

#### 4.17.4.2 Daunorubicin

Daunorubicin intercalates into DNA, blocks DNA-directed RNA polymerase and inhibits DNA synthesis. It can prevent cell division in doses that do not interfere with nucleic acid synthesis.

In combination with other drugs it is included in the first-choice chemotherapy of acute myelocytic leukemia in adults (for induction of remission), acute lymphocytic leukemia and the acute phase of chronic myelocytic leukemia. Oral absorption is poor, and it must be given intravenously. The half-life of distribution is 45 min and of elimination, about 19 h. The half-life of its active metabolite, daunorubicinol, is about 27 h. Daunorubicin is metabolized mostly in the liver and also secreted into the bile (ca 40%). Dosage must be reduced in liver or renal insufficiencies.



Suitable doses are (base equivalent), intravenous adult, younger than 60 yr. 45 mg/m<sup>2</sup>/day (30 mg/m<sup>2</sup> for patients older than 60 yr.) for 1, 2 or 3 days every 3 or 4 wk or 0.8 mg/kg/day for 3 to 6 days every 3 or 4 wk; no more than 550 mg/m<sup>2</sup> should be given in a lifetime, except only 450 mg/m<sup>2</sup> if there has been chest irradiation; children, 25 mg/m<sup>2</sup> once a  
5 wk unless the age is less than 2 yr. or the body surface less than 0.5 m, in which case the weight-based adult schedule is used. It is available in injectable dosage forms (base equivalent) 20 mg (as the base equivalent to 21.4 mg of the hydrochloride). Exemplary doses may be 10 mg/m<sup>2</sup>, 20 mg/m<sup>2</sup>, 30 mg/m<sup>2</sup>, 50 mg/m<sup>2</sup>, 100 mg/m<sup>2</sup>, 150 mg/m<sup>2</sup>, 175 mg/m<sup>2</sup>, 200 mg/m<sup>2</sup>, 225 mg/m<sup>2</sup>, 250 mg/m<sup>2</sup>, 275 mg/m<sup>2</sup>, 300 mg/m<sup>2</sup>, 350 mg/m<sup>2</sup>, 400 mg/m<sup>2</sup>, 425 mg/m<sup>2</sup>, 450 mg/m<sup>2</sup>,  
10 475 mg/m<sup>2</sup>, 500 mg/m<sup>2</sup>. Of course, all of these dosages are exemplary, and any dosage in-between these points is also expected to be of use in the invention.

#### 4.17.4.3 Mitomycin

Mitomycin (also known as mutamycin and/or mitomycin-C) is an antibiotic shown to  
15 have antitumor activity. Mitomycin selectively inhibits the synthesis of deoxyribonucleic acid (DNA). The guanine and cytosine content correlates with the degree of mitomycin-induced cross-linking. At high concentrations of the drug, cellular RNA and protein synthesis are also suppressed.

In humans, mitomycin is rapidly cleared from the serum after intravenous  
20 administration. Time required to reduce the serum concentration by 50% after a 30 mg. bolus injection is 17 min. After injection of 30 mg., 20 mg., or 10 mg. I.V., the maximal serum concentrations were 2.4 mg./ml, 1.7 mg./ml, and 0.52 mg./ml, respectively. Clearance is effected primarily by metabolism in the liver, but metabolism occurs in other tissues as well. The rate of clearance is inversely proportional to the maximal serum concentration because of  
25 saturation of the degradative pathways.

#### 4.17.4.4 Actinomycin D

Actinomycin D (Dactinomycin) is an antineoplastic drug that inhibits DNA-dependent  
RNA polymerase. It is a component of first-choice combinations for treatment of  
30 choriocarcinoma, embryonal rhabdomyosarcoma, testicular tumor and Wilms' tumor. Tumors that fail to respond to systemic treatment sometimes respond to local perfusion. Dactinomycin potentiates radiotherapy: It is a secondary (efferent) immunosuppressive.

Actinomycin D is used in combination with primary surgery, radiotherapy, and other drugs, particularly vincristine and cyclophosphamide. Antineoplastic activity has also been noted in Ewing's tumor, Kaposi's sarcoma, and soft-tissue sarcomas. Dactinomycin can be effective in women with advanced cases of choriocarcinoma. It also produces consistent responses in combination with chlorambucil and methotrexate in patients with metastatic testicular carcinomas. A response may sometimes be observed in patients with Hodgkin's disease and non-Hodgkin's lymphomas. Dactinomycin has also been used to inhibit immunological responses, particularly the rejection of renal transplants.

Half of the dose is excreted intact into the bile and 10% into the urine; the half-life is about 36 h. The drug does not pass the blood-brain barrier. Actinomycin D is supplied as a lyophilized powder (0.5 mg in each vial). The usual daily dose is 10 to 15 mg/kg; this is given intravenously for 5 days; if no manifestations of toxicity are encountered, additional courses may be given at intervals of 3 to 4 weeks. Daily injections of 100 to 400 mg have been given to children for 10 to 14 days; in other regimens, 3 to 6 mg/kg, for a total of 125 mg/kg, and weekly maintenance doses of 7.5 mg/kg have been used. Although it is safer to administer the drug into the tubing of an intravenous infusion, direct intravenous injections have been given, with the precaution of discarding the needle used to withdraw the drug from the vial in order to avoid subcutaneous reaction.

#### 4.17.4.5 Bleomycin

Bleomycin is a mixture of cytotoxic antibiotics that is freely soluble in water. Although the exact mechanism of action of bleomycin is unknown, available evidence would seem to indicate that the main mode of action is the inhibition of DNA synthesis with some evidence of lesser inhibition of RNA and protein synthesis.

In mice, high concentrations of bleomycin are found in the skin, lungs, kidneys, peritoneum, and lymphatics. Tumor cells of the skin and lungs have been found to have high concentrations of bleomycin in contrast to the low concentrations found in hematopoietic tissue. The low concentrations of bleomycin found in bone marrow may be related to high levels of bleomycin degradative enzymes found in that tissue.

In patients with a creatinine clearance of >35 ml per min, the serum or plasma terminal elimination half-life of bleomycin is approximately 115 min. In patients with a creatinine clearance of <35 ml per min, the plasma or serum terminal elimination half-life increases

exponentially as the creatinine clearance decreases. In humans, 60% to 70% of an administered dose is recovered in the urine as active bleomycin.

Bleomycin should be considered a palliative treatment. It has been shown to be useful in the management of the following neoplasms either as a single agent or in proven combinations with other approved chemotherapeutic agents in squamous cell carcinoma such as head and neck (including mouth, tongue, tonsil, nasopharynx, oropharynx, sinus, palate, lip, buccal mucosa, gingiva, epiglottis, larynx), skin, penis, cervix, and vulva. It has also been used in the treatment of lymphomas and testicular carcinoma.

Because of the possibility of an anaphylactoid reaction, lymphoma patients should be treated with two units or less for the first two doses. If no acute reaction occurs, then the regular dosage schedule may be followed.

Improvement of Hodgkin's Disease and testicular tumors is prompt and noted within 2 weeks. If no improvement is seen by this time, improvement is unlikely. Squamous cell cancers respond more slowly, sometimes requiring as long as 3 weeks before any improvement is noted.

#### 4.17.4.6 Cisplatin

Cisplatin has been widely used to treat cancers such as metastatic testicular or ovarian carcinoma, advanced bladder cancer, head or neck cancer, cervical cancer, lung cancer or other tumors. Cisplatin can be used alone or in combination with other agents, with efficacious doses used in clinical applications of 15-20 mg/m<sup>2</sup> for 5 days every three weeks for a total of three courses. Exemplary doses may be 0.50 mg/m<sup>2</sup>, 1.0 mg/m<sup>2</sup>, 1.50 mg/m<sup>2</sup>, 1.75 mg/m<sup>2</sup>, 2.0 mg/m<sup>2</sup>, 3.0 mg/m<sup>2</sup>, 4.0 mg/m<sup>2</sup>, 5.0 mg/m<sup>2</sup>, 10 mg/m<sup>2</sup>. Of course, all of these dosages are exemplary, and any dosage in-between these points is also expected to be of use in the invention.

Cisplatin is not absorbed orally and must therefore be delivered *via* injection intravenously, subcutaneously, intratumorally or intraperitoneally.

#### 4.17.4.7 VP16

VP16 is also known as etoposide and is used primarily for treatment of testicular tumors, in combination with bleomycin and cisplatin, and in combination with cisplatin for small-cell carcinoma of the lung. It is also active against non-Hodgkin's lymphomas, acute nonlymphocytic leukemia, carcinoma of the breast, and Kaposi's sarcoma associated with acquired immunodeficiency syndrome (AIDS).

VP16 is available as a solution (20 mg/ml) for intravenous administration and as 50-mg, liquid-filled capsules for oral use. For small-cell carcinoma of the lung, the intravenous dose (in combination therapy) is can be as much as  $100 \text{ mg/m}^2$  or as little as  $2 \text{ mg/m}^2$ , routinely  $35 \text{ mg/m}^2$ , daily for 4 days, to  $50 \text{ mg/m}^2$ , daily for 5 days have also been used. When given orally, the dose should be doubled. Hence the doses for small cell lung carcinoma may be as high as 200-250  $\text{mg/m}^2$ . The intravenous dose for testicular cancer (in combination therapy) is 50 to 100  $\text{mg/m}^2$  daily for 5 days, or 100  $\text{mg/m}^2$  on alternate days, for three doses. Cycles of therapy are usually repeated every 3 to 4 weeks. The drug should be administered slowly during a 30 to 60 min infusion in order to avoid hypotension and bronchospasm, which are probably due to the solvents used in the formulation.

#### 4.17.4.8 Tumor Necrosis Factor

Tumor Necrosis Factor [TNF; Cachectin] is a glycoprotein that kills some kinds of cancer cells, activates cytokine production, activates macrophages and endothelial cells, promotes the production of collagen and collagenases, is an inflammatory mediator and also a mediator of septic shock, and promotes catabolism, fever and sleep. Some infectious agents cause tumor regression through the stimulation of TNF production. TNF can be quite toxic when used alone in effective doses, so that the optimal regimens probably will use it in lower doses in combination with other drugs. Its immunosuppressive actions are potentiated by gamma-interferon, so that the combination potentially is dangerous. A hybrid of TNF and interferon- $\alpha$  also has been found to possess anti-cancer activity.

#### 4.17.4.9 Taxol and Taxotere

Taxol and taxotere are experimental antimitotic agents. They bind to tubulin and promote the assembly of microtubules. Taxol and taxotere have activity against malignant melanoma and carcinoma of the ovary. Maximal doses are  $30 \text{ mg/m}^2$  per day for 5 days or 210 to  $250 \text{ mg/m}^2$  given once every 3 weeks. Of course, all of these dosages are exemplary, and any dosage in-between these points is also expected to be of use in the invention.

#### 4.17.4.10 Vincristine

Vincristine blocks mitosis and produces metaphase arrest. It seems likely that most of the biological activities of this drug can be explained by its ability to bind specifically to tubulin and to block the ability of protein to polymerize into microtubules. Through disruption

of the microtubules of the mitotic apparatus, cell division is arrested in metaphase. The inability to segregate chromosomes correctly during mitosis presumably leads to cell death.

The relatively low toxicity of vincristine for normal marrow cells and epithelial cells make this agent unusual among anti-neoplastic drugs, and it is often included in combination  
5 with other myelosuppressive agents.

Unpredictable absorption has been reported after oral administration of vinblastine or vincristine. At the usual clinical doses the peak concentration of each drug in plasma is approximately 0.4 mM. Vinblastine and vincristine bind to plasma proteins. They are extensively concentrated in platelets and to a lesser extent in leukocytes and erythrocytes.

10 Vincristine has a multiphasic pattern of clearance from the plasma; the terminal half-life is about 24 h. The drug is metabolized in the liver, but no biologically active derivatives have been identified. Doses should be reduced in patients with hepatic dysfunction. At least a 50% reduction in dosage is indicated if the concentration of bilirubin in plasma is greater than 3 mg/dl (about 50 mM).

15 Vincristine sulfate is available as a solution (1 mg/ml) for intravenous injection. Vincristine used together with corticosteroids is presently the treatment of choice to induce remissions in childhood leukemia; the optimal dosages for these drugs appear to be vincristine, intravenously, 2 mg/m<sup>2</sup> of body-surface area, weekly, and prednisone, orally, 40 mg/m<sup>2</sup>, daily. Adult patients with Hodgkin's disease or non-Hodgkin's lymphomas usually receive vincristine  
20 as a part of a complex protocol. When used in the MOPP regimen, the recommended dose of vincristine is 1.4 mg/m<sup>2</sup>. High doses of vincristine seem to be tolerated better by children with leukemia than by adults, who may experience severe neurological toxicity. Administration of the drug more frequently than every 7 days or at higher doses seems to increase the toxic manifestations without proportional improvement in the response rate. Precautions should also  
25 be used to avoid extravasation during intravenous administration of vincristine. Vincristine (and vinblastine) can be infused into the arterial blood supply of tumors in doses several times larger than those that can be administered intravenously with comparable toxicity.

Vincristine has been effective in Hodgkin's disease and other lymphomas. Although it appears to be somewhat less beneficial than vinblastine when used alone in Hodgkin's disease,  
30 when used with mechlorethamine, prednisone, and procarbazine (the so-called MOPP regimen), it is the preferred treatment for the advanced stages (III and IV) of this disease. In non-Hodgkin's lymphomas, vincristine is an important agent, particularly when used with cyclophosphamide, bleomycin, doxorubicin, and prednisone. Vincristine is more useful than

vinblastine in lymphocytic leukemia. Beneficial response have been reported in patients with a variety of other neoplasms, particularly Wilms' tumor, neuroblastoma, brain tumors, rhabdomyosarcoma, and carcinomas of the breast, bladder, and the male and female reproductive systems.

5       Doses of vincristine for use will be determined by the clinician according to the individual patients need. 0.01 to 0.03 mg/kg or 0.4 to 1.4 mg/m<sup>2</sup> can be administered or 1.5 to 2 mg/m<sup>2</sup> can also be administered. Alternatively 0.02 mg/m<sup>2</sup>, 0.05 mg/m<sup>2</sup>, 0.06 mg/m<sup>2</sup>, 0.07 mg/m<sup>2</sup>, 0.08 mg/m<sup>2</sup>, 0.1 mg/m<sup>2</sup>, 0.12 mg/m<sup>2</sup>, 0.14 mg/m<sup>2</sup>, 0.15 mg/m<sup>2</sup>, 0.2 mg/m<sup>2</sup>, 0.25 mg/m<sup>2</sup> can be given as a constant intravenous infusion. Of course, all of these dosages are exemplary,  
10       and any dosage in-between these points is also expected to be of use in the invention.

#### 4.17.4.11     Vinblastine

When cells are incubated with vinblastine, dissolution of the microtubules occurs. Unpredictable absorption has been reported after oral administration of vinblastine or  
15       vincristine. At the usual clinical doses the peak concentration of each drug in plasma is approximately 0.4 mM. Vinblastine and vincristine bind to plasma proteins. They are extensively concentrated in platelets and to a lesser extent in leukocytes and erythrocytes.

After intravenous injection, vinblastine has a multiphasic pattern of clearance from the plasma; after distribution, drug disappears from plasma with half-lives of approximately 1 and  
20       20 h.

Vinblastine is metabolized in the liver to biologically active derivative desacetylvinblastine. Approximately 15% of an administered dose is detected intact in the urine, and about 10% is recovered in the feces after biliary excretion. Doses should be reduced in patients with hepatic dysfunction. At least a 50% reduction in dosage is indicated if the  
25       concentration of bilirubin in plasma is greater than 3 mg/dl (about 50 mM).

Vinblastine sulfate is available in preparations for injection. The drug is given intravenously; special precautions must be taken against subcutaneous extravasation, since this may cause painful irritation and ulceration. The drug should not be injected into an extremity with impaired circulation. After a single dose of 0.3 mg/kg of body weight, myelosuppression  
30       reaches its maximum in 7 to 10 days. If a moderate level of leukopenia (approximately 3000 cells/mm<sup>3</sup>) is not attained, the weekly dose may be increased gradually by increments of 0.05 mg/kg of body weight. In regimens designed to cure testicular cancer, vinblastine is used in doses of 0.3 mg/kg every 3 weeks irrespective of blood cell counts or toxicity.

The most important clinical use of vinblastine is with bleomycin and cisplatin in the curative therapy of metastatic testicular tumors. Beneficial responses have been reported in various lymphomas, particularly Hodgkin's disease, where significant improvement may be noted in 50 to 90% of cases. The effectiveness of vinblastine in a high proportion of  
5 lymphomas is not diminished when the disease is refractory to alkylating agents. It is also active in Kaposi's sarcoma, neuroblastoma, and Letterer-Siwe disease (histiocytosis X), as well as in carcinoma of the breast and choriocarcinoma in women.

Doses of vinblastine for use will be determined by the clinician according to the individual patients need. 0.1 to 0.3 mg/kg can be administered or 1.5 to 2 mg/m<sup>2</sup> can also be  
10 administered. Alternatively, 0.1 mg/m<sup>2</sup>, 0.12 mg/m<sup>2</sup>, 0.14 mg/m<sup>2</sup>, 0.15 mg/m<sup>2</sup>, 0.2 mg/m<sup>2</sup>, 0.25 mg/m<sup>2</sup>, 0.5 mg/m<sup>2</sup>, 1.0 mg/m<sup>2</sup>, 1.2 mg/m<sup>2</sup>, 1.4 mg/m<sup>2</sup>, 1.5 mg/m<sup>2</sup>, 2.0 mg/m<sup>2</sup>, 2.5 mg/m<sup>2</sup>, 5.0 mg/m<sup>2</sup>, 6 mg/m<sup>2</sup>, 8 mg/m<sup>2</sup>, 9 mg/m<sup>2</sup>, 10 mg/m<sup>2</sup>, 20 mg/m<sup>2</sup>, can be given. Of course, all of these dosages are exemplary, and any dosage in-between these points is also expected to be of use in the invention.

15

#### 4.17.4.12 Carmustine

Carmustine is one of the nitrosoureas used in the treatment of certain neoplastic diseases. It is highly soluble in alcohol and lipids, and poorly soluble in water. Carmustine is administered by intravenous infusion after reconstitution as recommended. Sterile carmustine  
20 is commonly available in 100 mg single dose vials of lyophilized material.

Although it is generally agreed that carmustine alkylates DNA and RNA, it is not cross resistant with other alkylators. As with other nitrosoureas, it may also inhibit several key enzymatic processes by carbamoylation of amino acids in proteins.

Carmustine is indicated as palliative therapy as a single agent or in established  
25 combination therapy with other approved chemotherapeutic agents in brain tumors such as glioblastoma, brainstem glioma, medulloblastoma, astrocytoma, ependymoma, and metastatic brain tumors. Also it has been used in combination with prednisone to treat multiple myeloma. Carmustine has proved useful, in the treatment of Hodgkin's Disease and in non-Hodgkin's lymphomas, as secondary therapy in combination with other approved drugs in patients who  
30 relapse while being treated with primary therapy, or who fail to respond to primary therapy.

The recommended dose of carmustine as a single agent in previously untreated patients is 150 to 200 mg/m<sup>2</sup> intravenously every 6 weeks. This may be given as a single dose or divided into daily injections such as 75 to 100 mg/m<sup>2</sup> on 2 successive days. When carmustine

is used in combination with other myelosuppressive drugs or in patients in whom bone marrow reserve is depleted, the doses should be adjusted accordingly. Doses subsequent to the initial dose should be adjusted according to the hematologic response of the patient to the preceding dose. It is of course understood that other doses may be used in the present invention for example 10 mg/m<sup>2</sup>, 20 mg/m<sup>2</sup>, 30 mg/m<sup>2</sup>, 40 mg/m<sup>2</sup>, 50 mg/m<sup>2</sup>, 60 mg/m<sup>2</sup>, 70 mg/m<sup>2</sup>, 80 mg/m<sup>2</sup>, 90 mg/m<sup>2</sup>, or 100 mg/m<sup>2</sup>. The skilled artisan is directed to, "Remington's Pharmaceutical Sciences" 15th Edition, chapter 61. Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject

10

#### 4.17.4.13 Melphalan

Melphalan is a bifunctional alkylating agent that is active against selective human neoplastic diseases. It is known chemically as 4-[bis(2-chloroethyl)amino]-L-phenylalanine.

Melphalan is the active L-isomer of the compound and was first synthesized in 1953 by Bergel and Stock; the D-isomer, known as medphalan, is less active against certain animal tumors, and the dose needed to produce effects on chromosomes is larger than that required with the L-isomer. The racemic (DL-) form is known as merphalan or sarcolysin. Melphalan is insoluble in water and has a pK<sub>a1</sub> of ~2.1. Melphalan is available in tablet form for oral administration and has been used to treat multiple myeloma.

Melphalan has been used in the treatment of epithelial ovarian carcinoma. One commonly employed regimen for the treatment of ovarian carcinoma has been to administer melphalan at a dose of 0.2 mg/kg daily for five days as a single course. Courses are repeated every four to five weeks depending upon hematologic tolerance (Young *et al.*, 1978). Alternatively the dose of melphalan used could be as low as 0.05 mg/kg/day or as high as 3 mg/kg/day or any dose in between these doses or above these doses. Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject

#### 30 4.17.4.14 Cyclophosphamide

Cyclophosphamide is stable in the gastrointestinal tract, tolerated well and effective by the oral and parental routes and does not cause local vesication, necrosis, phlebitis or even pain.



Suitable doses for adults include, orally, 1 to 5 mg/kg/day (usually in combination), depending upon gastrointestinal tolerance; or 1 to 2 mg/kg/day; intravenously, initially 40 to 50 mg/kg in divided doses over a period of 2 to 5 days or 10 to 15 mg/kg every 7 to 10 days or 3 to 5 mg/kg twice a wk or 1.5 to 3 mg/kg/day . A dose 250 mg/kg/day may be administered as an antineoplastic. Because of gastrointestinal adverse effects, the intravenous route is preferred for loading. During maintenance, a leukocyte count of 3000 to 4000/mm<sup>3</sup> usually is desired. The drug also sometimes is administered intramuscularly, by infiltration or into body cavities. It is available in dosage forms for injection of 100, 200 and 500 mg, and tablets of 25 and 50 mg the skilled artisan is referred to "Remington's Pharmaceutical Sciences" 15th Edition, chapter 61, incorporate herein as a reference, for details on doses for administration.

#### 4.17.4.15 Chlorambucil

Chlorambucil is a bifunctional alkylating agent of the nitrogen mustard type that has been found active against selected human neoplastic diseases. Chlorambucil is available in tablet form for oral administration. It is rapidly and completely absorbed from the gastrointestinal tract. After single oral doses of 0.6-1.2 mg/kg, peak plasma chlorambucil levels are reached within one h and the terminal half-life of the parent drug is estimated at 1.5 h. 0.1 to 0.2 mg/kg/day or 3 to 6 mg/m<sup>2</sup>/day or alternatively 0.4 mg/kg may be used for antineoplastic treatment. Treatment regimes are well know to those of skill in the art and can be found in the "Physicians Desk Reference" and in "Remingtons Pharmaceutical Sciences" referenced herein.

Chlorambucil is indicated in the treatment of chronic lymphatic (lymphocytic) leukemia, malignant lymphomas including lymphosarcoma, giant follicular lymphoma and Hodgkin's disease. It is not curative in any of these disorders but may produce clinically useful palliation.

#### 4.17.4.16 Busulfan

Busulfan is a bifunctional alkylating agent. Busulfan is available in tablet form for oral administration. Each scored tablet contains 2 mg busulfan and the inactive ingredients magnesium stearate and sodium chloride.

Busulfan is indicated for the palliative treatment of chronic myelogenous (myeloid, myelocytic, granulocytic) leukemia. Although not curative, busulfan reduces the total granulocyte mass, relieves symptoms of the disease, and improves the clinical state of the

patient. Approximately 90% of adults with previously untreated chronic myelogenous leukemia will obtain hematologic remission with regression or stabilization of organomegaly following the use of busulfan. It has been shown to be superior to splenic irradiation with respect to survival times and maintenance of hemoglobin levels, and to be equivalent to  
5 irradiation at controlling splenomegaly.

#### 4.17.4.17 Lomustine

Lomustine is one of the nitrosoureas used in the treatment of certain neoplastic diseases. Lomustine is soluble in 10% ethanol (0.05 mg per ml) and in absolute alcohol (70 mg per ml).  
10 Lomustine is relatively insoluble in water (<0.05 mg per ml). It is relatively unionized at a physiological pH. Although it is generally agreed that lomustine alkylates DNA and RNA, it is not cross resistant with other alkylators. As with other nitrosoureas, it may also inhibit several key enzymatic processes by carbamylation of amino acids in proteins.

Lomustine may be given orally. Following oral administration of radioactive lomustine  
15 at doses ranging from 30 mg/m<sup>2</sup> to 100 mg/m<sup>2</sup>, about half of the radioactivity given was excreted in the form of degradation products within 24 h. The serum half-life of the metabolites ranges from 16 h to 2 days. Tissue levels are comparable to plasma levels at 15 min after intravenous administration.

Lomustine has been shown to be useful as a single agent in addition to other treatment  
20 modalities, or in established combination therapy with other approved chemotherapeutic agents in both primary and metastatic brain tumors, in patients who have already received appropriate surgical and/or radiotherapeutic procedures. It has also proved effective in secondary therapy against Hodgkin's Disease in combination with other approved drugs in patients who relapse while being treated with primary therapy, or who fail to respond to primary therapy.

25 The recommended dose of lomustine in adults and children as a single agent in previously untreated patients is 130 mg/m<sup>2</sup> as a single oral dose every 6 weeks. In individuals with compromised bone marrow function, the dose should be reduced to 100 mg/m<sup>2</sup> every 6 weeks. When lomustine is used in combination with other myelosuppressive drugs, the doses should be adjusted accordingly. It is understood that other doses may be used for example, 20  
30 mg/m<sup>2</sup> 30 mg/m<sup>2</sup>, 40 mg/m<sup>2</sup>, 50 mg/m<sup>2</sup>, 60 mg/m<sup>2</sup>, 70 mg/m<sup>2</sup>, 80 mg/m<sup>2</sup>, 90 mg/m<sup>2</sup>, 100 mg/m<sup>2</sup>, 120 mg/m<sup>2</sup> or any doses between these ranges as determined by the clinician to be necessary for the individual being treated

## 5.0 EXAMPLES

### 5.1 Example 1: Identification of Prostate Markers by Use of Southern Differential Hybridization

Prostate enriched cDNAs were purchased from Clontech. The cDNAs were PCR<sup>TM</sup> amplified using adapter primers attached to both ends of the cDNAs. The amplified cDNAs were cloned into a pGEM-T plasmid vector (Promega) by T-A cloning and transformed into an appropriate bacterial host to generate a cDNA library. Transformed cells were plated and a total of 200 colonies were randomly picked from the prostate enriched cDNA library. The cDNA inserts were amplified by PCR<sup>TM</sup> from the plasmid templates using T7 and SP6 primers.  $\beta$ -Actin and PSA inserts were also prepared as controls. The inserts were run on duplicate 2% agarose gels and blotted onto nylon membranes. <sup>32</sup>P labeled cDNA probes were prepared, separately, from normal prostate RNA and pools of RNAs from 11 other tissues (liver, pancreas, testis, thymus, brain, mammary gland, skeletal muscle, kidney, lung, small intestine and spleen). The two membranes (each having identical amounts of cDNA from individual clones) were hybridized with the two probes (normal prostate or pool of other tissues) separately. Clones that hybridized only to the prostate cDNA probe were identified as potential prostate specific genes.

### 5.2 Example 2: Northern Analysis of Prostate Markers

Northern analysis was performed to confirm that the genes were expressed specifically in prostate tissue. Total cell RNA was isolated from human tissue samples and Northern blots were prepared according to Sambrook *et al.* (1989). The cDNA clones that were identified as potential prostate specific genes were <sup>32</sup>P labeled as probes, and hybridized against the Northern blots. UC Clone #51 (prostate-specific transglutaminase) message was preferentially expressed in prostate tissue relative to spleen, thymus, testis, ovary, small intestine, colon, and peripheral blood. UC Clone #57 (semenogelin II) message was preferentially expressed in normal prostate tissue when compared to the lack of detectable expression in samples taken from spleen, thymus, testis, ovary, small intestine, colon, and peripheral blood.

### 5.3 Example 3: DNA Sequences of Prostate Markers

The nucleotide sequences of prostate expressed clones were determined by dideoxy termination sequencing using either the ABI or Pharmacia automated sequencers. The DNA sequence of UC Clone #51 (SEQ ID NO:1) was identical in sequence to the sequence of prostate-specific transglutaminase (GenBank Accession #s L34840, I20492). The DNA sequence of Clone #56 (SEQ ID NO:2) was identical in sequence to the sequence of cytokeratin 15 (GenBank Accession # X07696). A third prostate specific gene, UC Clone #57 (SEQ ID NO:3) was identical in sequence to the sequence of semenogelin II (GenBank Accession # M81652). The identified sequences are provided in Table 5.

**TABLE 5. DNA Sequences of Prostate Markers**

UC Clone #51 (SEQ ID NO:1) prostate-specific transglutaminase, GenBank Accession #s L34840, I20492

```

5' AATTCTAAAAATGCTTTTGCAAGCTTGCATGCCTGCAGGTGCAGCGGCCGCCAGT
15 GTGATGGATATCTGCAGAATTCGGCTTGCGCTCAGCTGGAATTCGCAGAGATAGA
   GTCTTCCCTGGCATTGCAGGAGAGAATCTGAAGGGATGATGGATGCATCAAAAGA
   GCTGCAAGTTCTCCACATTGACTTCTTGAATCAGGACAACGCCGTTTCTCACCACA
   CATGGGAGTTCCAAACGAGCAGTCCTGTGTTCCGGCGAGGACAGGTGTTTCACCTG
   CGGCTGGTGCTGAACCAGCCCCTACAATCCTACCACCAACTGAACTGGAATTCAG
20 CACAGGGCCGAATCCTAGCATCGCCAAACACACCCTGGTGGTGCTCGACCCGAGG
   ACGCCCTCAGACCACTACAACCTGGCAGGCAACCCTTCAAAATGAGTCTGGCAAAG
   AGGTCACAGTGGCTGTCACCAGTTCCCCCAATGCCATCCTGGGCAAGTACCAACTA
   AACGTGAAAACCTGGAACACATCCTTAAGTCTGAAGAAAACATCCTATACCTTCT
   CTTCAACCCATGGTGTAAGAGGACATGGTTTTTCATGCCTGATGAGGACGAGCGCA
25 AAGAGTACATCCTCAATGACACGGGCTGCCATTACGTGGGGGCTGCCAGAAGTAT
   CAAATGCAAACCTGGAACCTTTGGTCAGTTTGAGAAAAATGTCCTGGACTGCTGCA
   TTTCCCTGCTGACTGAGAGCTCCCTCAAGCCACAGATAGGAGGGACCCCGTGCTG
   GTGTGCAGGGCCATGTGTGCTATGATGAGCTTTGAGAAAGGCCAGGGCGTGCTCAT
   TGGGAATTGGACTGGGGACTACGAAGGTGGCACAGCCCCATACAAGTGGACAGGC
30 AGTGCCCCGATCCTGCAGCAGTACTACAACACGAAGCAGGCTGTGTGCTTTGGCCA
   GTGCTGGGTGTTTGCTGGGATCCTGACTACAGTGCTGAGAGCGTTGGGCATCCCAG
   CACGCAGTGTGACAGGCTTCGATTGAGCTCACGACACAGAAAGGAACCTCACGGT
   GGACACCTATGTGAATGAGAATGGCGAGAAAATCACCAGTATGACCCACGACTCT
   GTCTGGAATTTCCATGTGTGGACGGATGCCTGGATGAAGCGACCCTACGACGGCTG
35 GCAGGCTGTGGACGCAACGCCGAGGAGCGAAGCCAGGGTGTCTTCTGCTGTGGG
   CCATCACCCTGACCGCCATCCGCAAAGGTGACATCTTTATTGTCTATGACACCAG
   ATTCGTCTTCTCAGAAGTGAATGGTGACAGGCTCATCTGGTTGGTGAAGATGGTGA
   ATGGGCAGGAGGAGTTACACGTAATTTCAATGGAGACCACAAGCATCGGGAAAAA
   CATCAGCACCAAGGCAGTGGGCCAAGACAGGCGGAGAGATATCACCTATGAGTAC
40 AAGTATCCAGAAGGCTCCTCTGAGGAGAGGCAGGTCATGGATCATGCCTTCCTCCT
   TCTCAGTTCTGAGAGGGAGCACAGACAGCCTGTAAAAGAGAACTTTCTTCACATGT
   CGGTACAATCAGATGATGTGCTGCTGGGAAACTCTGTTAATTTACCCGTGATTCTT

```

Table 5 – Continued

AAAAGGAAGACCGCTGCCCTACAGAATGTCAACATCTTGGGCTCCTTTGAACTACA  
GTTGTACACTGGCAAGAAGATGGCAAACTGTGTGACCTCAATAAGACCTCGCAG  
5 ATCCAAGGTCAAGTATCAGAAGTGACTCTGACCTTGGACTCCAAGACCTACATCAA  
CAGCCTGGCTATATTAGATGATGAGCCAGTTATCAGAGGTTTCATCATTGCGGAAA  
TTGTGGAGTCTAAGGAAATCATGGCCTCTGAAGTATTCACGTCAAACCAGTACCCT  
GAGTTCTCTATAGAGTTGCCTAACACAGGCAGAATTGGCCAGCTACTTGTCTGCAA  
TTGTATCTTCAAGAATACCCTGGCCATCCCTTTGACTGACGTCAAGTTCTCTTTGGA  
10 AAGCCTGGGCATCTCCTCACTACAGACCTCTGACCATGGGACGGTGCAGCCTGGTG  
AGACCATCCAATCCCAAATAAAATGCACCCCAATAAΛΛΛCTGGACCCAAAGAAATT  
TATCGTCAAGTTAAGTTCCAAACAAGTGAAAGAGATTAATGCTCAGAAGATTGTTC  
TCATCACCAAGTAGCCTTGTCTGATGCTGTGGAGCCTTAGTTGAGATTTACAGCATTT  
CCTACCTTGTGCTTAGCTTTCAGATTATGGATGATTAAATTTGATGACTTATATGAG  
15 GGCAGATTCAAGAGCCAGCAGGTCAAAAAGGCCAACACAACCATAAGCAGCCAG  
ACCCACAAGGCCAGGTCTGTGCTATCACAGGGTACCTCTTTACAGTTAGAAAC  
ACCAGCCGAGGCCACAGAATCCCATCCCTTTCTGAGTCATGGCCTCAAAAATCAG  
GGCCACCATTGTCTCAATTCAAATCCATAGATTTCGAAGCCACAGAGCTCTTCCCT  
GGAGCAGCAGACTATGGGCAGCCAGTGCTGCCACCTGCTGACGACCCTTGAGAA  
20 GCTGCCATATCTTCAGGCCATGGGTTCACCAGCCCTGAAGGCACCTGTCAACTGGA  
GTGCTCTCTCAGCACTGGGATGGGCCTGATAGAAGTGCATTCTCCTCCTATTGCCTC  
CATTCTCCTCTCTATCCCTGAAATCCAGGAAGTCCCTCTCCTGGTGCTCCAAGCA  
GTTTGAAGCCCAATCTGCAAGGACATTTCTCAAGGGCCATGTGGTTTTGCAGACAA  
CCCTGTCTCAGGCCTGAACTCACCATAGAGACCCATGTCAGCAAACGGTGACCAG  
25 CAAATCCTCTTCCCTTATTCTAAAGCTGCCCTTGGGAGACTCCAGGGAGAAGGCA  
TTGCTTCTCCTGGTGTGAACTCTTTCTTTGGTATTCCATCCACTATCCTGGCAACT  
CAAGGCTGCTTCTGTAACTGAAGCCTGCTCCTTCTTGTCTGCCCTCCAGAGATT  
GCTCAAATGATCAATAAGCTTTAAATTAACCGGAATCCGCGGAATTC-3'

30

UC Clone #56 (SEQ ID NO:2) cytokeratin 15, GenBank Accession # X07696

5'GGTACCTCCTGCCAGCACCTCTTGGGTTTGCTGAGAACTCACGGGCTCCAGCTAC  
CTGGCCATGACCACCACATTTCTGCAAACCTTCTCCTCCACCTTTGGGGGTGGCTCA  
35 ACCCGAGGGGGTTCCCTCCTGGCTGGGGGAGGTGGCTTTGGTGGGGGGAGTCTCTC  
TGGGGGAGGTGGAAGCCGAAGTATCTCAGCTTCTTCTGCTAGGTTTGTCTCTCAG  
GGTCAGGAGGAGGATATGGGGGTGGCATGAGGGTCTGTGGCTTTGGTGGAGGGGC  
TGGTAGTGTTTTCGGTGGAGGCTTTGGAGGGGGCGTTGGTGGGGGTTTTGGTGGTG  
GCTTTGGTGGTGGCGATGGTGGTCTCCTCTCTGGCAATGAGAAAATTACCATGCAG  
40 AACCTCAATGACCGCCTGGCCTCCTACCTGGACAAGGTACGTGCCCTGGAGGAGGC  
CAATGCTGACCTGGAGGTGAAGATCCATGACTGGTACCAGAAGCAGACCCAGCC  
AGCCAGAAATGCGACTACAGCCAATACTTCAAGACCATTGAAGAGCTCCGGGACA  
AGATCATGGCCACCACCATCGACAACCTCCCGGGTCATCCTGGAGATCGACAATGCC  
AGGCTGGCTGCGGACGACTTCAGGCTCAAGTATGAGAATGAGCTGGCCCTGCGCC  
45 AGGGCGTTGAGGCTGACATCAACGGCTTGCGCCGAGTCCTGGATGAGCTGACCCTG  
GCCAGGACTGACCTGGAGATGCAGATCGAGGGCCTGAATGAGGAGCTAGCCTACC  
TGAAGAAGAACCACGAAGAGGAGATGAAGGAGTTCAGCAGCCAGCTGGCCGGCC  
AGGTCAATGTGGAGATGGACGCAGCACCGGGTGTGGACCTGACCCGTGTGCTGGC

Table 5 – Continued

AGAGATGAGGGAGCAGTACGAGGCCATGGCGGAGAAGAACCGCCGGGATGTCGA  
 GGCCTGGTTCTTCAGCAAGACTGAGGAGCTGAACAAAGAGGTGGCCTCCAACACA  
 5 GAAATGATCCAGACCAGCAAGACGGAGATCACAGACCTGAGACGCACGATGCAGG  
 AGCTGGAGATCGAGCTGCAGTCCCAGCTCAGCATGAAAGCTGGGCTGGAGAACTC  
 ACTGGCCGAGACAGAGTGCCGCTATGCCACGCAGCTGCAGCAGATCCAGGGGCTC  
 ATTGGTGGCCTGGAGGGCCAGCTGAGTGAGCTCCGATGCGAGATGGAGGCTCAGA  
 ACCAGGAGTACAAGATGCTGCTTGACATAAAGACACGGCTGGAGCAGGAGATCGC  
 10 TACTTACCGCAGCCTGCTCGAGGGCCAGGATGCCAAGATGGCTGGCATTGGCATCA  
 GGGAAGCCTCTTCAGGAGGTGGTGGTAGCAGCAGCAATTTCCACATCAATGTAGA  
 AGAGTCAGTGGATGGACAGGTGGTTTCTTCCCACAAGAGAGAAATCTAAGTGTCTA  
 TTGCAGGAGAAACGTCCCTTGCCACTCCCCACTCTCATCAGGCCAAGTGGAGGACT  
 GGCCAGAGGGCCTGCACATGCAAACCTCCAGTCCCTGCCTTCAGAGAGCTGAAAAG  
 15 GGTCCCTCGGTCTTTTATTTTCAGGGCTTTGCATGCGCTCTATTCCCCCTCTGCCTCTC  
 CCCACCTTCTTTGGAGCAAGGAGATGCAGCTGTATTGTGTAACAAGCTCATTTGTA  
 CAGTGTCTGTTTCATGTAATAAAGAATTACTTTTCCTTTTGCAAAT-3'

20 UC Clone #57 (SEQ ID NO:3) semenogelin II, GenBank Accession # M81652

5'AGACAAGATTTTTCAAGCAAGATGAAGTCCATCATCCTCTTTGTCTTTCCCTGCT'  
 CCTTATCTTGGAGAAGCAAGCAGCTGTGATGGGACAAAAAGGTGGATCAAAAGGC  
 CAATTGCCAAGCGGATCTTCCCAATTTCCACATGGACAAAAGGGCCAGCACTATTT  
 25 TGGACAAAAAGACCAACAACATACTAAATCCAAAGGCAGTTTTTCTATTCAACACA  
 CATATCATGTAGACATCAATGATCATGACTGGACCCGAAAAAGTCAGCAATATGAT  
 TTGAATGCCCTACATAAGGCGACAAAATCAAAACAACACCTAGGTGGAAGTCAAC  
 AACTGCTCAATTATAAACAAGAAGGCAGAGACCATGATAAATCAAAAGGTCATTT  
 TCACATGATAGTTATACATCATAAAGGAGGCCAAGCTCATCATGGGACACAAAAT  
 30 CCTTCTCAAGATCAGGGGAATAGCCCATCTGGAAAGGGATTATCCAGTCAATGTTT  
 AAACACAGAAAAAAGGCTATGGGTTTCATGGACTAAGTAAAGAACAGCTTCAGCC  
 TCTGGTGCACAAAAAGGTAGAACACAAGGTGGATCCCAAGCAGTTATGTTCTCC  
 AAACTGAAGAACTAGTAGTTAACAACAACAACGTGAGACTAAAAATTCTCATCA  
 AAATAAAGGGCATTACCAAAATGTGGTTGACGTGAGAGAGGAACATTCAAGTAAA  
 35 CTACAAACTTCACTCCATCCTGCACATCAAGACAGACTCCAACATGGACCCAAAGA  
 CATTTTTACTACCAAGATGAGCTCCTAGTATATAACAAGAATCAACACCAGACAA  
 AAAATCTCAGTCAAGATCAAGAGCATGGCCGGAAGGCACATAAAATATCATACCC  
 GTCTTCACGTACAGAAGAAAGACAACCTTCACCATGGAGAAAAGAGTGTACAGAAA  
 GATGTATCCAAAGGCAGCATTCTATCCAAACTGAAGAGAAAAATACATGGCAAGT  
 40 CTCAAAACCAGGTAACAATTCATAGTCAAGATCAAGAGCATGGCCATAAGGAAAA  
 TAAAATATCATACCAATCTTCAAGTACAGAAGAAAGACATCTCAACTGTGGAGAA  
 AAGGGCATCCAGAAAGGTGTATCCAAAGGCAGTATTTTCGATCCAAACTGAAGAGC  
 AAATACATGGCAAGTCTCAAAACCAGGTAAGAATTCCTAGTCAAGCTCAAGAGTA  
 TGGCCATAAGGAAAAATAAAATATCATACCAATCTTCGAGTACAGAAGAAAGACGT  
 45 CTCAACAGTGGAGAAAAGGATGTACAGAAAGGTGTATCCAAAGGCAGTATTTCTA  
 TCCAAACTGAAGAGAAAATACATGGCAAGTCTCAAAACCAGGTAACAATTCCTAG  
 TCAAGATCAAGAGCATGGCCATAAGGAAAAATAAAATGTCATACCAATCTTCAAGT  
 ACAGAAGAAAGACGACTCAACTATGGAGGAAAGAGCACGCAGAAAGATGTATCC

Table 5 – Continued

CAAAGCAGTATTTCTTTCCAAATTGAAAAGCTAGTAGAAGGCAAGTCTCAAATCCA  
GACACCAAATCCTAATCAAGATCAATGGTCTGGCCAAAATGCAAAAAGGAAAGTCT  
5 GGTCAATCTGCAGATAGCAAACAAGACCTACTCAGTCATGAACAAAAAGGCAGAT  
ACAAACAGGAATCCAGTGAGTCACATAATATTGTAATTACTGAGCATGAGGTTGCC  
CAAGATGATCATTTGACACAACAATATAATGAAGACAGAAATCCAATATCTACAT  
AGCCCTGTTGCTTAGCAACCACTTGAAAAGCTGGACCAATAGCAAGGTGTCACCCG  
ACCTCAGTGAAGTCTTTGATGTTTCTGAGAGGCAGACTCCCATGTGGTCCCAGATC  
10 CTTGGTCCATGGATGACACCACCTTCCCATGCTTCCTTGCATTAGGCTTTCTAAACC  
CGGAGCCCCTTCAAACCTCCAATAAAGGGATCATTCTGCTTT-3'

#### 5.4 Example 4: Relative Quantitative Reverse Transcriptase-Polymerase Chain Reaction

15 Relative quantitative RT-PCR™ was used to independently confirm the differential expression of the mRNAs identified above. The reverse transcription-polymerase chain reaction (RT-PCR™) protocol used in the following examples was described in US Application Serial No. 08/692,787, incorporated in relevant part herein by reference. Total cell RNA was  
20 converted into cDNA using reverse transcriptase primed with random hexamers. Preferred methods for RNA isolation are the guanidinium thiocyanate method and kits for RNA isolation manufactured by Qiagen, Inc. (Chatworth, CA).

The RNAs were digested with DNase I to remove all genomic DNA. Prior to DNase I digestion, the RNA was in a particulate suspension in 70% ethanol. Approximately 50 µg of  
25 RNA (as determined by OD<sub>260/280</sub>) was removed from the suspension, precipitated and resuspended in DEPC treated sterile water. To this was added 10X DNase I buffer (200 mM Tris-HCl; pH 8.4, 20 mM MgCl<sub>2</sub>, 500 mM KCl), 10 units of RNase Inhibitor (GIBCO-BRL Cat#15518-012) and 20 units of DNase I (GIBCO-BRL # 18068-015). The volume was adjusted to 50 µl with additional DEPC treated water and incubated at 37°C for 30 min. After  
30 DNase I digestion the RNAs were extracted with phenol and chloroform followed by ethanol precipitation.

The total cell RNAs were analyzed using a PCR™-based test to confirm that the DNase I treatment digested the contaminating genomic DNA to completion. The assay for contaminating genomic DNA utilized oligonucleotides that flank a 145 nucleotide long intron  
35 (intron #3) in the gene encoding Prostate Specific Antigen (PSA). This is a single copy gene with no pseudogenes. The oligonucleotides used in this assay are specific to PSA. The sequences of these oligonucleotides were:

5'CGCCTCAGGCTGGGGCAGCATT 3', SEQ ID NO:4

and

5'ACAGTGGAAGAGTCTCATTCGAGAT 3', SEQ ID NO:5.

5

Aliquots of 500 ng to 1.0 µg of each of the DNase I treated RNAs were used as templates in a standard PCR™ (35-40 cycles). Human genomic DNA was used as a positive control, resulting in the amplification of a 242 nucleotide PCR™ product visualized on an ethidium bromide stained electrophoretic gel.

10

The standard conditions used for PCR™ were: 1X GIBCO-BRL PCR™ reaction buffer [20 mM Tris-Cl (pH 8.4), 50 mM KCl], 1.5 mM MgCl<sub>2</sub>, 200 µM each of the four dNTPs, 200 nM each oligonucleotide primer, template as appropriate and 2.5 units of Taq polymerase per 100 µl of reaction volume. Using these conditions, PCR™ was performed with 35-40 cycles of 94°C for 45 sec, 55°-60°C for 45 sec, and 72°C for 1 min.

15

Reverse transcription reactions were performed using the Superscript™ Preamplification System for First Strand cDNA Synthesis kit from GIBCO-BRL LifeTechnologies (Gaithersburg, MD), using the manufacturer's protocols. As template, either 5 or 10 micrograms of RNA was used. The RT reaction product was diluted with water to a final volume of 100 µl.

20

cDNAs from total cell RNAs were normalized to contain equal concentrations of amplifiable β-actin cDNA. One µl of each diluted RT reaction was subjected to PCR™ using oligonucleotides specific to β-actin as primers. These β-actin specific oligonucleotides have the sequences:

25

5' CGAGCTGCCTGACGGCCAGGTCATC 3', SEQ ID NO:6

and

5' GAAGCATTTGCGGTGGACGATGGAG 3', SEQ ID NO:7

30

PCR™ was performed under standard conditions for either 19 or 20 cycles, resulting in a 415 bp PCR™ product. The product was examined by agarose gel electrophoresis followed by staining with ethidium bromide and visualized by irradiation with a transilluminator. An image of the illuminated gel was captured by an IS-1000 Digital Imaging System manufactured



by Alpha Innotech Corporation. The captured image was analyzed using either version 2.0 or 2.01 of the software package supplied by the manufacturer to determine the relative amounts of amplified  $\beta$ -actin cDNA in each RT reaction. To normalize the various cDNAs, water was added to the most concentrated cDNAs. PCR<sup>TM</sup> using 1  $\mu$ l of the newly rediluted and adjusted cDNA is repeated using the  $\beta$ -actin primers. With this empirical method the cDNAs were adjusted by dilution to contain roughly equal concentrations of amplifiable cDNA.

PCR<sup>TM</sup> was performed using different gene specific oligonucleotides as primers to determine the relative abundances of other mRNAs as represented as cDNAs in the normalized panel of diluted RT reaction products. The relative intensities of the bands were adjusted and normalized to  $\beta$ -actin.

For the genes isolated in this study, total cell RNA was isolated from metastatic prostate cancer or buffy coat cells as described above. cDNA was made from one to five  $\mu$ g of each isolated RNA. All cDNAs were normalized for similar amounts of  $\beta$ -actin cDNA by RT-PCR<sup>TM</sup>.

For relative quantitative RT-PCR<sup>TM</sup> with an external standard, quantitation of band intensities on ethidium bromide stained gels was performed using the IS-1000 image analysis system. A normalizing statistic was generated for each cDNA sample, as the average of all  $\beta$ -actin signals divided by the  $\beta$ -actin signal for each cDNA sample respectively. Data for each sample was then normalized by multiplying the observed densitometry observation by the individual normalizing statistics. Normalized values predict differences in the steady state abundances of the respective mRNAs in the original total cell RNA samples.

This protocol resulted in the discovery that the expression of two cDNAs, UC Clone #51 (SEQ ID NO:1), UC Clone #56 (SEQ ID NO:2), was down regulated in metastatic prostate cancer, and the expression of one cDNA, UC Clone #57 (SEQ ID NO:3), was down regulated in the peripheral blood of metastatic prostate cancer patients.

UC Clone #51 (SEQ ID NO:1) was confirmed by relative quantitative RT-PCR<sup>TM</sup> to be down regulated in metastatic prostate cancer tissues in comparison to normal prostate and organ confined prostate cancer, including BPH. The data was normalized against  $\beta$ -actin mRNA. This gene was down-regulated to the point of its expression being totally inhibited in metastatic cancer patients when compared with normal and BPH individuals. Such a clear contrast in regulation makes this gene an excellent marker for the detection of malignant prostate tumors in biopsy samples containing a mixture of normal, benign and malignant prostate cells.

UC Clone #56 (SEQ ID NO:2) was confirmed by relative quantitative RT-PCR<sup>TM</sup> to be down regulated in metastatic prostate cancer tissues in comparison to normal prostate and organ

confined prostate cancer, including BPH. The data was normalized against  $\beta$ -actin mRNA. This gene was down-regulated in metastatic cancer patients compared with normal and BPH individuals, making it a useful marker for metastatic prostate cancer.

UC Clone #57 (SEQ ID NO:3) was not differentially regulated in prostate cancer tissues compared to normal prostate. However, relative quantitative RT-PCR™ of UC Clone #57 (semenogelin II) determined that expression of this gene was down regulated in the blood of individuals with metastatic prostate cancer compared to normal individuals. Those who are skilled in the art will recognize the usefulness of a metastatic prostate marker that can be easily obtained from peripheral blood, as opposed to collection from a prostate tissue biopsy.

Of the genes quantitated with these primers, prostate-specific transglutaminase (GenBank Accession #s L34840, I20492) and cytokeratin 15 (GenBank Accession # X07696) were more abundant in normal and BPH glands and are potential tumor suppressor genes. Semenogelin II (GenBank Accession # M81652 and M81651) was more abundant in the peripheral blood of patients with metastatic prostate cancer and is contemplated to be a progression marker.

The oligonucleotides used for relative quantitative RT-PCR™ are listed in Table 6. These sequences are designated herein as SEQ ID NO:8 (GenBank Accession #s L34840, I20492, prostate-specific transglutaminase Nt 548-571); SEQ ID NO:9 (GenBank Accession #s L34840, I20492, prostate-specific transglutaminase Nt 742-765, antisense strand); SEQ ID NO:10 (GenBank Accession # X07696, cytokeratin 15 Nt 1337-1359); SEQ ID NO:11 (GenBank Accession # X07696, cytokeratin 15 Nt 1586-1608, antisense strand); SEQ ID NO:12 (GenBank Accession # M81652 semenogelin II Nt 1089-1116); SEQ ID NO:13 (GenBank Accession # M81652, semenogelin II Nt 1697-1724, antisense strand).

**TABLE 6. Oligonucleotides used in relative quantitative RT-PCR™.**

Oligonucleotides used to examine the expression of genes:

Prostate-specific transglutaminase (SEQ ID NO:1), GenBank Accession #L34840, I20492.

5' GGGGGCTGCCAGAAGTATCAAATG3', SEQ ID NO:8

5' TGCCACCTTCGTAGTCCCCAGTCC3', SEQ ID NO:9

Cytokeratin 15 (SEQ ID NO:2), GenBank Accession #X07696.

5' TCTTCAGGAGGTGGTGGTAGCAG3', SEQ ID NO:10

Table 6 – Continued

5' GAGAGGCAGAGGGGGAATAGAGC3', SEQ ID NO:11  
 Semenogelin II (SEQ ID NO:3), GenBank Accession #M81652 and M81651.  
 5' ACATCTCAACTGTGGAGAAAAGGGCATC3', SEQ ID NO:12  
 5' TGATCATCTTGGGCAACCTCATGCTCAG3', SEQ ID NO:13  
 Controls used to normalize relative quantitative RT-PCR™  
 Prostate Specific Antigen (PSA)  
 5' CGCCTCAGGCTGGGGCAGCATT 3', SEQ ID NO:4  
 5' ACAGTGGAAGAGTCTCATTCGAGAT 3', SEQ ID NO:5.

β-actin

5' CGAGCTGCCTGACGGCCAGGTCATC3', SEQ ID NO:6

5' GAAGCATTTGCGGTGGACGATGGAG3', SEQ ID NO:7

A summary of experiments performed to confirm the aforementioned genes as prostate disease markers are shown below in Table 7.

TABLE 7

Genes Whose mRNAs have Abundances that Vary in Prostate Disease Relative to Normal Individuals

Name of cDNA Fragment	Sequence Determined	Confirmed by Northern analysis	Confirmed by RT-PCR™	Previously Known
UC Clone #51 (SEQ ID NO:1)	Yes	Yes	Yes	GB #L34840, GB #I20492
UC Clone #56 (SEQ ID NO:2)	Yes	No	Yes	GB #X07696
UC Clone #57 (SEQ ID NO:3)	Yes	Yes	Yes	GB #M81652

It will be recognized that the genes and gene products (RNAs and proteins) for these markers of prostate disease are included within the scope of the disclosure herein described. It will also be recognized that the diagnosis and prognosis of prostatic disease by detection of the nucleic acid products of these genes are included within the scope of the present invention. Serological and other assays to detect these mRNA species or their translation products are also indicated. It is

obvious that these assays are of utility in diagnosing metastatic cancers derived from prostate and other tissues.

Those practiced in the art will realize that there exists naturally occurring genetic variation between individuals. As a result, some individuals may synthesize prostate-specific transglutaminase, cytokeratin 15, or semenogelin II gene products that differ from those described by the sequences entailed in the Genbank number listed above. Included in the definition of prostate-specific transglutaminase, cytokeratin 15, or semenogelin II are those products encoded by prostate-specific transglutaminase, cytokeratin 15, or semenogelin II genes that vary in sequence from those described above. Those practiced in the art will realize that modest variations in DNA sequence will not significantly obscure the identity of a gene product as being derived from the prostate-specific transglutaminase, cytokeratin 15, or semenogelin genes.

#### 5.5 Example 5: Prostate specific Transglutaminase Promoter

The prostate specific transglutaminase gene possesses two features that make it specifically useful in the treatment of prostate disease: 1) prostate specific expression; and 2) dramatic down regulation in high Gleason grade and metastatic prostate cancer. Cloning of the promoter for this gene would allow the analysis and isolation of regulatory regions that provide its tissue-specific expression. Promoter availability would also facilitate the development of methods for inhibition or up-regulation of gene expression in metastatic prostate cancer.

To clone the promoter for prostate specific transglutaminase, a human placental genomic DNA library in the EMBL phage vector was obtained from Clontech (Palo Alto, CA) and screened. The probe used for screening was a 217 bp PCR<sup>TM</sup> product amplified from prostate specific transglutaminase cDNA using the A3-P1 and A3-P2 primers:

A3-P1: 5' GGG GGC TGC CAG AAG TAT CAA ATG 3' (SEQ ID NO:16)

A3-P2: 5' TGC CAC CTT CGT AGT CCC CAG TCC 3' (SEQ ID NO:17)

Eight 150-mm LB agar plates were used for the primary screening. For each plate, 10 µl of diluted phage (about 50,000 pfu) were combined with 0.6 ml host cells in a 15-ml tube for each plate. The tubes were incubated at 37°C for 15 min before addition of 9 ml of melted LB soft agarose. The contents of each tube were poured onto separate LB agar plates. The plates were inverted and incubated at 37°C for 6-8 h until the plaques reached a diameter not

exceeding 2 mm or were just beginning to make contact with one another. The plates were then chilled at 4°C overnight.

Nylon membranes were marked and placed in direct contact with the top of each plate. After 1-2 min, the membrane was carefully peeled off and immersed in DNA denaturing solution (0.5 N NaOH, 1.5 NaCl) for 3 min. The membrane was then immersed in neutralizing solution (0.5 M Tris-Cl, pH 8.0, 1.5 M NaCl) for 3 min. Nucleic acids were cross-linked to the membrane using a UV crosslinker (Stratalinker, from Stratagene, La Jolla, CA).

The membranes were prehybridized in 10 ml of hybridization solution (RapidHyb, Amersham, UK) at 68°C for 15 min with agitation. The 217 bp probe, prepared as described above, was labeled with <sup>32</sup>P-α-dCTP using random-primed labeling (Amersham, UK). After the labeled DNA probe was denatured by heating at 100°C for 10 min and rapid chilling on ice, it was added to the pre-hybridization solution (10<sup>6</sup> cpm/ml). Hybridization was performed at 68°C for 2 h. After hybridization, the membrane was washed at room temperature in wash solution #1 (2 X SSC, 0.1% SDS) for 20 min, in wash solution #2 (1 X SSC, 0.1% SDS) twice at 65°C for 30 min, and in wash solution #3 (0.1 X SSC, 0.1% SDS) for 20 min. The membranes were exposed to X-ray film at -70°C for 3 days using an intensifying screen to identify positive clones.

Two positive plaques were identified from the initial screening. About 500 plaques from each positive candidate were re-plated and rescreened as described above. Following secondary screening, a single clone (identified as A3GM1) with a 1.45 kb DNA insert was obtained. The insert of this clone was PCR™ amplified using primers flanking the cloning site and subcloned into the pGEM-T plasmid vector.

The insert was sequenced by dideoxy termination sequencing. The full sequence of this genomic clone is shown in FIG 1. The 3' terminal 53 bp sequence of the genomic clone matches the 5' end of the human prostate specific transglutaminase cDNA (GenBank Accession #L34840, SEQ ID NO:1), confirming that this genomic clone corresponds to the promoter region of the gene. The complete sequence of this clone, including the 5' end of the prostate specific transglutaminase gene, is shown in FIG. 1. A listing of the promoter sequence, along with the 5' end of the transcribed region of the gene, is provided in SEQ ID NO:15.

The TRANSFAC database was searched for potential regulatory elements present in the prostate specific transglutaminase promoter using two computer based software programs, Signal Scan (bimas.dcrn.nih.gov/molbio/signal) and MatInspector

(www.gsf.de/cgi-bin/matsearch). Of course, other software programs and databases known to those of skill in the art may be used to search for putative regulatory or transcription factor binding elements, such as, for example, the Genetics Computer Group program (GSG, Madison, WI). Table 8 below is an example of how putative regulatory sites may be identified through comparison to known regulatory sequences. Over 50 possible transcriptional factors binding sites and responsive elements were identified (Table 8). Sequences identified in this Table or through the techniques described herein or known to one of skill in the art may be targeted for mutation, alteration, duplication, truncation, or combination with other regulatory sequences from other genes using any of the techniques described herein or known to one of skill in the art. Additional putative elements may be discovered through comparison to the regulatory regions of other gene sequences and through comparison with other known consensus or putative gene regulatory sites.

**Table 8. Possible regulatory sites in Prostate specific Transglutaminase Promoter**

Sites	Locations (from transcription start site)	Sequences
GATA1	-1369	CagtGATAtgctc
LMO2COM	-1367	GtGATAtgc
CEBPB	-1356	tcattgtGTAAata
XFD1	-1352	tgtGTAAataaac
SRY	-1344	taaaACAAata
HFH2	-1329	catTATTtgctt
HNF3B	-1329	catTATTtgctt
RFX1	-1299	caggagccctGTAAcat
CMYB	-1254	caacgtcaccGTTGgcgg
ATF	-1252	ACGTCA
NF1	-1245	cgtTGGCggcaagcctt
S8	-1119	ttcacagaATTatcc
API	-1084	ggTGACTcat
MZF1	-1042	tatGGGGa
AP2	-1040	TGGGGA
BRN2	-982	tccatggcAAATgctg
SP1	-934	TGCAC

Table 8 - Continued

SRF	-963	CCAAT
PADS	-899	tGTGGTctt
CEBPB	-895	gtctttgCAAaggt
NF1	-893	cttTGGCaaggtgaagggc
AP1	-806	gcTGACtgtgg
AP4	-796	gaCAGCtctt
GATA1	-768	TTATCT
MZF1	-735	tcaGGGGa
CAC binding	-720	CCACC
GF11	-701	taacctagAATCttgtgctctcg
CEBPB	-604	tgattatGAAAata
OCT1	-603	gattatgaAAATaaa
SRY	-582	aataACAAaata
HNF3B	-558	catTATTtgcct
GR	-536	ACAACA
BRN2	-439	tatattcaAAATgtaa
TFIID	-435	TTCAAA
c-Myc	-404	TCTCTTA
GATA1	-321	agcttGATAgtaac
GATA1	-320	gccttGATAgtaac
LMO2COM	-318	ttGATAgta
PADS	-288	tGTGGTgta
RORA1	-286	tggtgtaGGTCat
CMYB	-219	gccattgtccGTTGtctg
GR	-203	TGTCCC
VBP	-196	tTTACatcag
AP4_	-191	atCAGCtcat
NF1	-182	tctTGGCtcttgattcc
GATA1	-162	ggcttGATAtggaa
GATA1	-161	gtctGATAtggaa

Table 8 – Continued

LMO2COM	-159	ctGATAtgg
IK2	-146	ttccGGGAtttt
BRN2	-134	attatttcTAATcaac
Sp1	-105	ACCCGCCCC
Tjian GC Box	-104	CCCGCCC
NF1	-97	cctTGGCagccatagcaa
CP2	-89	gccatagCAAG
TATA(Goldbery-Hogness)	-42	TATAA

### 5.6 Example 6: Prostate Specific Expression

5 The 1.4 kb promoter, the 5' end of the promoter (-520 to -1400) and the 3' end of the promoter (-1 to -520) were cloned into pCAT3 reporter vectors. The promoter showed prostate specific activity and the 5' end of the promoter was responsible for prostate specific expression of the gene. The 1.4 kb fragment was cloned into pCAT3 (Promega) at Mlu I and Xho I sites (FIG. 2). The promoter linked to a CAT reporter gene or a control plasmid were transfected into prostate cancer cell line PC3 and bladder cancer cell line T24 using Lipofectamine (Life Technologies). A CAT activity assay was performed using the FAST CAT Green kit from Molecular Probes (Eugene, OR). Both PC3 and T24 cells transfected with the positive control plasmid (pCAT) showed CAT activity (FIG. 2). However, only PC3 cells showed CAT activity when the prostate specific transglutaminase promoter was used to drive CAT expression (FIG. 2). No CAT activity was seen in T24 cells with this promoter (FIG. 2).

The 5' end of the promoter (-520 to -1400) was cloned into pCAT3 at the Mlu I and Bgl II sites to test its activity (FIG. 3). As shown in FIG. 3, both PC3 and T24 cells showed CAT activity when transfected with the pCAT3 control plasmid. However, when transfected with the 5' end of the transglutaminase promoter, PC3 cells showed much higher activity than cells transfected with the control pCAT3 plasmid, indicating that the 5' end of the transglutaminase promoter is much more effective at transcription than the SV40 promoter in prostate cancer cells (compare FIG. 2 *versus* FIG. 3). The T24 bladder cancer line showed no CAT expression with the 5' end of the transglutaminase promoter (FIG. 3). This demonstrates that elements in the 5' end of the promoter regulate and enhance prostate specific expression of the gene and that elements in the 5' end of the promoter inhibit gene expression in cells other than prostate.



**REFERENCES**

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

- 5 EPO 239,400
- EPO 273,085
- EPO 320,308
- EPO 329,822
- PCT US 87/00880
- 10 PCT US 89/01025
- GB 2,202,328
- U.S. Patent Application Serial No. 08/692,787
- U.S. Patent No. 3,817,837
- U.S. Patent No. 3,850,752
- 15 U.S. Patent No. 3,939,350
- U.S. Patent No. 3,996,345
- U.S. Patent No. 4,196,265
- U.S. Patent No. 4,237,224
- U.S. Patent No. 4,275,149
- 20 U.S. Patent No. 4,277,437
- U.S. Patent No. 4,366,241
- U.S. Patent No. 4,578,770
- U.S. Patent No. 4,596,792
- U.S. Patent No. 4,599,230
- 25 U.S. Patent No. 4,599,231
- U.S. Patent No. 4,601,903
- U.S. Patent No. 4,608,251
- U.S. Patent No. 4,642,234
- U.S. Patent No. 4,683,195
- 30 U.S. Patent No. 4,683,202
- U.S. Patent No. 4,797,368
- U.S. Patent No. 4,800,159
- U.S. Patent No. 4,883,750

- U.S. Patent No. 4,946,773  
U.S. Patent No. 5,139,941  
U.S. Patent No. 5,168,053  
U.S. Patent No. 5,279,721  
5 U.S. Patent No. 5,354,855  
U.S. Patent No. 5,527,695  
U.S. Patent No. 5,545,818  
U.S. Patent No. 5,599,672  
U.S. Patent No. 5,614,396  
10 U.S. Patent No. 5,624,824  
U.S. Patent No. 5,624,830  
U.S. Patent No. 5,625,047  
U.S. Patent No. 5,631,236  
U.S. Patent No. 5,658,772  
15 U.S. Patent No. 5,783,435  
U.S. Patent No. 5,783,681  
U.S. Patent No. 5,821,337  
U.S. Patent No. 5,830,876  
WO 88/10315.  
20 WO 89/06700.  
WO 90/07641  
WO 94/10343  
WO 97/41228  
Adler *et al.*, *M. Mol. Encocrinol.*, 5:1587-1596, 1991.  
25 Alder *et al.*, *Mol. Cell. Biol.*, 13:6326-6335, 1993.  
Amster *et al.*, *Nucleic Acids Res.* 8:2055, 1980  
Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press,  
Cold Spring Harbor, New York, 1988.  
Araki *et al.*, *J. Mol. Biol.* 225(1):25-37, 1992.  
30 Baichwal and Sugden, *In*: Kucherlapati R, ed. *Gene transfer*. New York: Plenum Press, pp.  
117-148, 1986.  
Baker *et al.*, *Science*, 249:912-915, 1990.  
Barry *et al.*, *Nature* 377: 632-635, 1995.

- Beato, *Cell*, 56:335-344, 1989.
- Bellus, *J. Macromol. Sci. Pure Appl. Chem*, A31(1): 1355-1376, 1994.
- Berzal-Herranz *et al.*, *Genes and Devel.*, 6:129-134, 1992.
- Bird *et al.*, *Science*, 242, 423-426, 1988
- 5 Bittner *et al.*, *Methods in Enzymol*, 153: 516-544, 1987.
- Boss *et al.*, *Nucleic Acids Res.* 12:3791, 1984
- Boyer and Weiner, *FASEB J.*, 12: 1611-1626, 1998.
- Brooks *et al.*, *Prostate*, 35: 18-26, 1998.
- Bussemakers *et al.*, *Proc. Am. Assoc. Cancer. Res.* 37:515, 1996.
- 10 Cabilly *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3273, 1984
- Call *et al.*, *Cell*, 60:509-520, 1990.
- Capaldi *et al.*, *Biochem. Biophys. Res. Comm.*, 76: 425, 1977.
- Cawthon *et al.*, *Cell*, 62:193-201, 1990.
- Cech *et al.*, *Cell*, 27:487-496, 1981.
- 15 Chandler *et al.*, *The Plant Cell*, 1:1175-1183, 1989.
- Chang *et al.*, *Hepatology*, 14:134A, 1991.
- Chau *et al.*, *Science*, 244:174-181, 1989.
- Chen and Okayama, *Mol. Cell. Biol.* 7:2745-2752, 1987.
- Chothia *et al.*, *J. Mol. Biol.* 196:901, 1987
- 20 Chowrira *et al.*, *Biochemistry*, 32:1088-1095, 1993.
- Chowrira *et al.*, *J. Biol. Chem.*, 269:25856-25864, 1994.
- Chung *et al.*, *Cancer Research*, 54:2577-2581, 1994.
- Clark *et al.*, *Human Gene Therapy*, 6:1329-1341, 1995.
- Cleutjens *et al.*, *Mol. Endocrinol.*, 11(9):1256-1265, 1997.
- 25 Coffin, *In: Virology*, Fields *et al.* (eds.), New York: Raven Press, pp. 1437-1500, 1990.
- Colberre-Garapin *et al.*, *J. Mol. Biol.*, 150: 1, 1981.
- Cole *et al.*, *Genomics* 51: 282-287, 1998.
- Cordon-Cardo *et al.*, *J. Natl. Cancer Inst.* 90: 1284-1291, 1998.
- Corr *et al.*, *J. Urol.* 151:492A, 1994.
- 30 Cotten *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:6094-6098, 1992.
- Couch *et al.*, *Am. Rev. Resp. Dis.*, 88:394-403, 1963.
- Coupar *et al.*, *Gene*, 68:1-10, 1988.

- Curiel, In: *Viruses in Human Gene Therapy*, J.-M.H. Vos (Ed.), Carolina Academic Press, Durham, N.C., pp. 179-212, 1994.
- Dai and Burnstein, *Molecular Endocrinology*, 10:1582-1594, 1996.
- Daugherty *et al.*, *Nucleic Acids Res.* 19:2471, 1991
- 5 Dong *et al.*, *Science*, 268:884-886, 1995.
- Duan *et al.*, *Science*, 269:1402-1406, 1995.
- Ebert *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:5745-5749, 1987.
- Ellis *et al.*, *EMBO Journal*, 6(11):3203-3208, 1987.
- Enomoto, *et al.*, *J Bacteriol.*, 6(2):663-668, 1983.
- 10 Fearon *et al.*, *Cell*, 61:759-767, 1990a.
- Fearon *et al.*, *Science*, 247:49-56, 1990b.
- Fechheimer *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:8463-8467, 1987.
- Finlay *et al.*, *Cell*, 57:1083-1093, 1989.
- Flotte *et al.*, *Am. J. Respir. Cell Mol. Biol.*, 7:349-356, 1992.
- 15 Flotte *et al.*, *Gene Therapy*, 2:29-37, 1995.
- Flotte *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:10613-10617, 1993.
- Forster and Symons, *Cell*, 49:211-220, 1987.
- Fraley and Kaplan, *Proc. Nat'l. Acad. Sci. USA*, 76:3348-3352, 1979.
- Freifelder, *Physical Biochemistry Applications to Biochemistry and Molecular Biology*,  
20 2nd ed., Wm. Freeman and Co., New York, NY, 1982.
- Friedmann, *Science*, 244:1275-1281, 1989.
- Friend *et al.*, *Nature*, 323:643-646, 1986.
- Frohman, *PCR™ Protocols: A Guide to Methods and Applications*, Academic Press, N.Y.,  
1990.
- 25 Fung *et al.*, *Science*, 236:1657-1660, 1987.
- Futreal *et al.*, *Science*, 266:120-122, 1994.
- Gelvin *et al.*, In: *Plant Molecular Biology Manual*, 1990.
- Gerlach *et al.*, *Nature (London)*, 328:802-805, 1987.
- Gessler *et al.*, *Nature*, 343:774-778, 1990.
- 30 Ghosh-Choudhury *et al.*, *EMBO J.*, 6:1733-1739, 1987.
- Ghosh and Bachhawat, In: Wu G. and C. Wu (eds.) *Liver diseases, targeted diagnosis and therapy using specific receptors and ligands*. New York: Marcel Dekker, pp. 87-104, 1991.

- Ghossein *et al.*, *J. Clin. Oncol.*, 13:1195-1200, 1995.
- Gillman and Smith, *Gene*, 8:81-97, 1979
- Goethuys *et al.*, *Am. J. Clin. Oncol.*, 20(1):40-45, 1997.
- Gomella *et al.*, *J. Urology*, 158:326-337, 1997.
- 5 Gomez-Foix *et al.*, *J. Biol. Chem.*, 267:25129-25134, 1992.
- Goodrich and Lee, *Biochim. Biophys. Acta*, 1155:43-61, 1993.
- Gopal, *Mol. Cell. Biol.*, 5:1188-1190, 1985.
- Gotoh *et al.*, *J. Urology*, 160:220-229, 1998.
- Graham and Prevec, *Biotechnology*, 20:363-390, 1992.
- 10 Graham and Prevec, *In: Gene Transfer and Expression Protocols*, Murray, E.J., ed., Humana, New Jersey, vol. 7, 109-128, 1991.
- Graham and Van Der Eb, *Virology* 52:456-467, 1973.
- Graham *et al.*, *J. Gen. Virol.*, 36:59-72, 1977.
- Grunhaus and Horwitz, *Seminar in Virology*, 3:237-252, 1992.
- 15 Harland and Weintraub, *J. Cell Biol.*, 101:1094-1099, 1985.
- Haseloff and Gerlach, *Nature*, 334:585-591, 1988.
- Haseloff *et al.*, *Proc Natl Acad Sci U S A*, 18;94(6):2122-2127, 1997.
- Hermonat and Muzyczka, *Proc. Nat'l. Acad. Sci. USA*. 81:6466-6470, 1984.
- Hersdorffer *et al.*, *DNA Cell Biol.*, 9:713-723, 1990.
- 20 Herz and Gerard, *Proc. Nat'l. Acad. Sci. USA*, 90:2812-2816, 1993.
- Horwich *et al.*, *J. Virol.*, 64:642-650, 1990.
- Hrouda and Dalglish, *Gene Ther.*, 3(10):845-852, 1996.
- Huang *et al.*, *Science*, 242:1563-1566, 1988.
- Hudspeth and Grula, *Plant Mol. Biol.*, 12:579-589, 1989.
- 25 Huston *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 85:5879-5883, 1988.
- Ikuta *et al.*, *Bio/technol.*, 8:241-242, 1990.
- Innis *et al.*, *PCR™ Protocols*, Academic Press, Inc., San Diego CA, 1990.
- Inouye *et al.*, *Nucleic Acids Res.*, 13:3101-3109, 1985.
- Israeli *et al.*, *Cancer Res.*, 53:227-230, 1993.
- 30 Israeli *et al.*, *Urol. Clin. North. Am.*, 24(2):439-450, 1997.
- Jang *et al.*, *J. Virology*, 63:1651-1660, 1989.
- Jefferson, *Plant Mol. Biol. Rep.*, 5:387-405, 1987.
- Jones and Shenk, *Cell*, 13:181-188, 1978.

- Joshi, *Nucleic Acids Res.*, 15:6643-6653, 1987.
- Joyce, *Nature*, 338:217-244, 1989.
- Kamb *et al.*, *Science*, 264:436-440, 1994.
- Kaneda *et al.*, *Science*, 243:375-378, 1989.
- 5 Kaplitt *et al.*, *Nature Genetics*, 8:148-154, 1994.
- Karlsson *et al.*, *EMBO J.*, 5:2377-2385, 1986.
- Kasahara *et al.*, *Science*, 266:1373-1376, 1994.
- Kato *et al.*, *J. Biol. Chem.*, 266:3361-3364, 1991.
- Katz *et al.*, *J. Gen. Microbiol.*, 129:2703-2714, 1983.
- 10 Kelleher and Vos, *Biotechniques*, 17(6):1110-1117, 1994.
- Kim and Cech, *Proc. Natl. Acad. Sci. USA*, 84:8788-8792, 1987.
- Klein *et al.*, *Nature*, 327:70-73, 1987.
- Koller and Smithies, 1992.
- Konety and Getzenberg, *Semin. Urol. Oncol.*, 15(1):33-42, 1997.
- 15 Kotin *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:2211-2215, 1990.
- Kwoh *et al.*, *Proc. Nat. Acad. Sci. USA*, 86:1173, 1989.
- LaFace *et al.*, *Virology*, 162:483-486, 1988.
- Landis *et al.*, *CA Cancer J. Clin.*, 48: 6-29, 1998.
- Lanzavecchia *et al.*, *Eur. J. Immunol.*, 17:105, 1987
- 20 Laughlin *et al.*, *J. Virol.*, 60:515-524, 1986.
- Lawton *et al.*, *Plant Mol. Biol.* 9:315-324, 1987.
- Le Gal La Salle *et al.*, *Science*, 259:988-990, 1993.
- Lebkowski *et al.*, *Mol. Cell. Biol.*, 8:3988-3996, 1988.
- Lee *et al.*, *Science*, 235:1394-1399, 1987a.
- 25 Lee *et al.*, *Nature*, 329:642-645, 1987b.
- Levrero *et al.*, *Gene*, 101:195-202, 1991.
- Lieber and Strauss, *Mol. Cell. Biol.*, 15:540-551, 1995.
- Liu *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:3439, 1987
- Lowy *et al.*, *Cell*, 22:817, 1980.
- 30 Luo *et al.*, *Blood* 82(Supp):1,303A, 1994.
- Ma *et al.*, *Nature*, 334:631-633, 1988.
- Macejak and Sarnow, *Nature*, 353:90-94, 1991.
- Maeser and Kahmann, *Mol. Gen. Genetics*, 230:170-176, 1991.

- Malkowicz and Johnson, *Gene Ther.*, 12:649-664, 1998.
- Mann *et al.*, *Cell*, 33:153-159, 1983.
- Markowitz *et al.*, *J. Virol.*, 62:1120-1124, 1988.
- Marshall, *Cell*, 64:313-326, 1991.
- 5 Martiniello-Wilks *et al.*, *Hum. Gene Ther.* 9(11):1617-1626, 1998.
- Massuda *et al.*, *Proc. Natl. Acad. Sci. USA.*, 94:14701-14706, 1997.
- Matteucci *et al.*, *J. Am. Chem. Soc.*, 103:3185, 1981.
- Mbikay *et al.*, *DNA*, 6: 23-29, 1987.
- McCarty *et al.*, *J. Virol.*, 65:2936-2945, 1991.
- 10 McLaughlin *et al.*, *J. Virol.*, 62:1963-1973, 1988.
- Michel and Westhof, *J. Mol. Biol.*, 216:585-610, 1990.
- Miki *et al.*, *Science*, 266:66-71, 1994.
- Miller, *Curr. Top. Microbiol. Immunol.*, 158:1, 1992.
- Mulligan *et al.*, *Proc. Nat'l Acad. Sci. USA*, 78:2072, 1981.
- 15 Murphy *et al.*, *Prostate* 38: 73-78, 1999.
- Murtha *et al.*, *Biochemistry* 32:6459-6464, 1993.
- Muzyczka, *Curr. Top. Microbiol. Immunol.*, 158:97-129, 1992.
- Nakamura *et al.*, In: *Enzyme Immunoassays: Heterogeneous and Homogeneous Systems*, Chapter 27, 1987b.
- 20 Nakamura *et al.*, In: *Handbook of Experimental Immunology* (4th Ed.), Weir, E., Herzenberg, L.A., Blackwell, C., Herzenberg, L. (eds), Vol. 1, Chapter 27, Blackwell Scientific Publ., Oxford, 1987a.
- Ng *et al.*, *Development*, 120:3257-3265, 1994.
- Nicolas and Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt (eds.), Stoneham: Butterworth, pp. 494-513, 1988.
- 25 Nicolau and Sene, *Biochim. Biophys. Acta*, 721:185-190, 1982
- Nicolau *et al.*, *Methods Enzymol.*, 149:157-176, 1987.
- O'Hare *et al.*, *Proc. Nat'l Acad. Sci. USA*, 78:1527, 1981.
- Odell *et al.*, *Nature*, 313:810-812, 1985.
- 30 Ohara *et al.*, *Proc. Nat'l Acad. Sci. USA*, 86:5673-5677, 1989.
- Ohi *et al.*, *Gene*, 89L:279-282, 1990.
- Olssen *et al.*, *Urologic Clinics of North America (USA)*, 24/2: 367-378, 1997.
- Ow *et al.*, *Science*, 234:856-859, 1986.

- Palukaitis *et al.*, *Virology*, 99:145-151, 1979.
- Paskind *et al.*, *Virology*, 67:242-248, 1975.
- Pelletier and Sonenberg, *Nature*, 334:320-325, 1988.
- Perales *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:4086-4090, 1994.
- 5 Perriman *et al.*, *Gene*, 113:157-163, 1992.
- Phi-Van *et al.*, *Mol. Cell. Biol.*, 10:2302-2307, 1990.
- Poszkowski *et al.*, *EMBO J.*, 3:2719, 1989.
- Potrykus *et al.*, *Mol. Gen. Genet.*, 199:183-188, 1985.
- Potter *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:7161-7165, 1984.
- 10 Prasher *et al.*, *Biochem. Biophys. Res. Commun.*, 126(3):1259-1268, 1985.
- Prescott *et al.*, *Prostate*, 35:71-80, 1998.
- Pritchard-Jones *et al.*, *Nature*, 346:194-197, 1990.
- Prody *et al.*, *Science*, 231:1577-1580, 1986.
- Racher *et al.*, *Biotechnology Techniques*, 9:169-174, 1995.
- 15 Ragot *et al.*, *Nature*, 361:647-650, 1993.
- Reichel *et al.*, *Proc. Natl. Acad. Sci. USA*, 93 (12):5888-5893, 1996.
- Reinhold-Hurek and Shub, *Nature*, 357:173-176, 1992.
- Reiter *et al.*, *Proc. Nat'l. Acad. Sci. USA*, 95:1735-1740, 1998.
- Remington's Pharmaceutical Sciences, 15<sup>th</sup> Ed., Mack Publishing Co., Easton, PA 1975.
- 20 Renan, *Radiother. Oncol.*, 19:197-218, 1990.
- Rich *et al.*, *Hum. Gene Ther.*, 4:461-476, 1993.
- Ridgeway, In: *Vectors: A survey of molecular cloning vectors and their uses*. Rodriguez and Denhardt, eds. Stoneham: Butterworth, pp. 467-492, 1988.
- Rippe *et al.*, *Mol. Cell Biol.*, 10:689-695, 1990.
- 25 Roberts *et al.*, *Nature*, 328:731-734, 1987.
- Robinson *et al.*, *Nature Medicine*, 5:526-534, 1999.
- Roman *et al.*, *Nature Medicine*, 3:849-854, 1997.
- Rosenfeld *et al.*, *Cell*, 68:143-155, 1992.
- Rosenfeld *et al.*, *Science*, 252:431-434, 1991.
- 30 Roux *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:9079-9083, 1989.
- Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989.
- Samulski *et al.*, *EMBO J.*, 10:3941-3950, 1991.



- Samulski *et al.*, *J. Virol.*, 63:3822-3828, 1989.
- Santerre *et al.*, *Gene*, 30:147, 1984.
- Sarver *et al.*, *Science*, 247:1222-1225, 1990.
- Scanlon *et al.*, *Proc Natl Acad Sci USA*, 88:10591-10595, 1991.
- 5 Scarlett and Robins, *Mol. Endocrinol.* 9, 413-423, 1995.
- Schott *et al.*, *Cancer Res.*, 54:1393-1396, 1994.
- Segawa *et al.*, *Cancer Res.* 58:2282-2287, 1998.
- Serrano *et al.*, *Nature*, 366:704-707, 1993.
- Shan *et al.*, *Endocrinology*, 138(9):3764-3770, 1997.
- 10 Sheen *et al.*, *Plant Journal*, 8(5):777-784, 1995.
- Shelling and Smith, *Gene Therapy*, 1:165-169, 1994.
- Shida *et al.*, *Mol. Endocrinol.* 7:181-188, 1993.
- Sioud *et al.*, *J Mol. Biol.*, 223:831-835, 1992.
- Stalker *et al.*, *Science*, 242:419-422, 1988.
- 15 Stief *et al.*, *Nature* 341:343, 1989.
- Stratford-Perricaudet *et al.*, *Hum. Gene. Ther.*, 1:241-256, 1991.
- Stratford-Perricaudet and Perricaudet, *In: Human Gene Transfer*. Eds. Cohen-Haguenaer and Boiron, Editions John Libbey Eurotext, France, pp. 51-61. 1991.
- Sullivan *et al.*, *Mol. Gen. Genet.*, 215(3):431-440, 1989.
- 20 Sutcliffe, *Proc. Natl. Acad. Sci. USA*, 75:3737-3741, 1978.
- Symons, *Annu. Rev. Biochem.*, 61:641-671, 1992.
- Symons, *Nucl. Acids Res.*, 9:6527-6537, 1981.
- Szybalska *et al.*, *Proc. Nat'l Acad. Sci. USA*, 48:2026, 1962.
- Takehiko *et al.*, *Cancer Res.*, 58:2282-2287, 1998.
- 25 Temin, *In: Gene Transfer*, Kucherlapati (ed.), New York: Plenum Press, pp. 149-188, 1986.
- Thillet *et al.*, *J. Biol. Chem.*, 263:12500-12508, 1988.
- Thompson *et al.*, *Nature Genet.*, 9:444-450, 1995.
- Tian, Sequin, Charest, *Plant Cell Rep.*, 16:267-271, 1997.
- Tilley *et al.*, *J. Biol. Chem.*, 265: 13776-781, 1990.
- 30 Tjoa *et al.*, *Prostate* 40: 125-129, 1999.
- Tomic *et al.*, *Nucl. Acids Res.*, 12:1656, 1990.
- Top *et al.*, *J. Infect. Dis.*, 124:155-160, 1971.
- Tratschin *et al.*, *Mol. Cell. Biol.*, 4:2072-2081, 1984.

- Tratschin *et al.*, *Mol. Cell. Biol.*, 5:32581-3260, 1985.
- Tur-Kaspa *et al.*, *Mol. Cell Biol.*, 6:716-718, 1986.
- Ulmer *et al.*, *Science*, 259:1745-1749, 1993.
- Upender, Raj, Weir, *Biotechniques*, 18(1):29-30, 1995.
- 5 Veltri *et al.*, *J. Urol.*, 155:551A, 1996.
- Vile and Hart, *Cancer Res.* 53:962-967, 1993.
- Viskochil *et al.*, *Cell*, 62:187-192, 1990.
- Wagner *et al.*, *Science*, 260:1510-1513, 1990.
- Walker *et al.*, *Proc. Nat'l Acad. Sci. USA*, 89:392-396, 1992.
- 10 Walker *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:6624-6628, 1987.
- Wall *et al.*, *Nucleic Acids Res.*, 5:3113, 1978.
- Wallace *et al.*, *Science*, 249:181-186, 1990.
- Walsh *et al.*, *J. Clin. Invest.*, 94:1440-1448, 1994.
- Wang *et al.*, *Molecular and Cellular Biology*, 12(8):3399-3406, 1992.
- 15 Ward *et al.*, *Nature*, 341:544, 1989.
- Wei *et al.*, *Gene Therapy*, 1:261-268, 1994.
- Weinberg, *Science*, 254:1138-1146, 1991.
- Weiner and Kennedy, *Sci. Amer.* 281: 50-57, 1999.
- Wigler *et al.*, *Cell*, 11:223, 1977.
- 20 Wigler *et al.*, *Proc. Nat'l Acad. Sci. USA*, 77:3567, 1980.
- Wong *et al.*, *Eur. J. Biochem.*, 221:917-925, 1994.
- Wong *et al.*, *Gene*, 10:87-94, 1980.
- Wu and Wu, *Adv. Drug Delivery Rev.*, 12:159-167, 1993.
- Wu and Wu, *J. Biol. Chem.*, 262:4429-4432, 1987.
- 25 Yan *et al.*, *Prostate*, 32(2):129-139, 1997.
- Yang and Russell, *Proc. Natl. Acad. Sci. USA*, 87:4144-4148, 1990.
- Yang *et al.*, *Proc. Nat'l Acad. Sci. USA*, 87:9568-9572, 1990.
- Yang *et al.*, *J. Virol.*, 68:4847-4856, 1994.
- Yang *et al.*, *J. Urol.*, 155:552A, 1996.
- 30 Yang *et al.*, *Clin. Cancer Res.* 4: 1873-1880, 1998.
- Yoder *et al.*, *Blood*, 82(Supp.):1:347A, 1994.
- Yoo *et al.*, *J. Biol. Chem.*, 764:17078, 1989
- Young *et al.*, *J. Androl.* 16:97-99, 1995.

- Yuan and Altman, *Science*, 263:1269-1273, 1994.
- Yuan *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:8006-8010, 1992.
- Zalsut *et al.*, *Nucleic Acids Res.*, 8:3591, 1980.
- Zelivianski *et al.*, *Biochem. Biophys. Res. Commun.*, 245(1):108-112, 1998.
- 5 Zhang *et al.*, *Nucleic Acid Research*, 25:3143-3150, 1997.
- Zhou *et al.*, *Exp. Hematol. (NY)*, 21:928-933, 1993.
- Zhou, *et al.*, *J. Exp. Med.*, 179:1867-1875, 1994.
- Zou *et al.*, *Science*, 263:526-529, 1994.
- Zukowsky *et al.*, *Proc. Natl. Acad. Sci. USA*, 80:1101-1105, 1983.

**CLAIMS:**

1. An isolated nucleic acid comprising between 20 and 1453 contiguous nucleotides of the sequence of SEQ ID NO:15.
- 5 2. An isolated nucleic acid comprising a prostate specific transglutaminase promoter.
3. The isolated nucleic acid of claim 2, wherein said promoter comprises the sequence of SEQ ID NO:15.
- 10 4. A prostate specific transglutaminase promoter isolatable from the nucleic acid sequence of SEQ ID NO:15.
5. An expression vector comprising a prostate specific transglutaminase promoter operably linked to a selected gene.
- 15 6. The vector of claim 5, wherein said promoter comprises the nucleic acid sequence of SEQ ID NO:15.
7. The vector of claim 5, wherein said gene encodes prostate specific transglutaminase,  
20 cytokeratin 15 or semenogelin II.
8. The vector of claim 5, wherein said gene encodes thymidine kinase, p53, cytosine deaminase, PNP, fibroblast growth factor receptor 2 (FGFR2), nitroreductase, PTEN, FHIT, KAI1 or diptheria toxin.
- 25 9. The vector of claim 5, wherein said gene encodes a cytokine.
10. The vector of claim 5, wherein said gene encodes an anti-bcl-2 ribozyme, an anti-sense c-myc RNA, E-cadherin, or polyglutamine.
- 30 11. The vector of claim 5, wherein said gene encodes hepatitis B surface antigen, herpes simplex glycoprotein, HIV envelope proteins, hemagglutinin, malaria circumsporozoite protein,

carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveollin, POV1, HER2/neu or p27KIP1.

12. The vector of claim 5, wherein said gene encodes a humanized antibody.

5

13. The vector of claim 12, wherein humanized antibody is Herceptin®.

14. The vector of claim 12, wherein said humanized antibody binds specifically to PSMA, PSCA, caveolin, POV1, HER2/neu or p27KIP1.

10

15. The vector of claim 6, wherein a portion of the sequence of SEQ ID NO:15 is deleted, said portion between 10 and 1350 nucleotides in length.

16. The vector of claim 5, wherein said promoter comprises between 20 and 1453

15 contiguous nucleotides of the sequence of SEQ ID NO:15.

17. The vector of claim 5, further comprising an enhancer.

18. The vector of claim 17, wherein said enhancer is the prostate specific antigen (PSA)

20 enhancer.

19. A composition comprising an isolated nucleic acid having the sequence of SEQ ID NO:15.

25 20. A composition comprising an isolated nucleic acid complementary to SEQ ID NO:15.

21. A genetic vaccine comprising a prostate specific transglutaminase promoter operably linked to a selected gene.

30 22. The genetic vaccine of claim 21, wherein said gene encodes hepatitis B surface antigen, herpes simplex glycoprotein, HIV envelope proteins, hemagglutinin, malaria circumsporozoite protein, carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveollin, POV1, HER2/neu or p27KIP1

23. A method of identifying a prostate specific promoter comprising:
- a) providing a nucleic acid probe of a sequence selected from prostate specific transglutaminase, cytokeratin 15 or semenogelin II;
  - 5 b) screening a human genomic library with said probe;
  - c) identifying a clone that hybridizes under high stringency conditions with said probe; and
  - d) confirming that said clone comprises a prostate specific promoter.
- 10 24. A method of identifying a prostate specific promoter comprising:
- a) providing a nucleic acid probe of a sequence identical to or fully complementary with SEQ ID NO:15;
  - b) screening a human genomic library with said probe;
  - c) identifying a clone that hybridizes under reduced stringency conditions with said probe; and
  - 15 d) confirming that said clone comprises a prostate specific promoter.
25. A method of identifying protein binding factors for a prostate specific promoter comprising:
- 20 a) providing an isolated, double-stranded nucleic acid molecule comprising the sequence of SEQ ID NO:15;
  - b) providing nuclei from cells of prostate origin;
  - c) extracting proteins from said nuclei;
  - d) allowing said proteins to bind specifically to said nucleic acid molecule;
  - 25 e) removing unbound proteins;
  - f) isolating proteins bound specifically to said nucleic acid molecule; and
  - g) identifying said proteins.
26. A method of identifying regulatory sequences within the promoter of prostate specific transglutaminase comprising:
- 30 a) providing an isolated, double-stranded nucleic acid comprising the sequence of SEQ ID NO:15;

- b) making at least one deletion mutant of said nucleic acid, wherein said deletion mutant is missing a portion of the sequence of SEQ ID NO:15, said portion between approximately 10 and 1350 basepairs in length;
- c) operably linking said deletion mutant to a reporter gene; and
- 5 d) assaying the amount of expression of said reporter gene linked to said deletion mutant;

wherein the presence of a regulatory sequence within the deleted portion of SEQ ID NO:15 is indicated by a change in the expression of said reporter gene, compared to the expression of said reporter gene operably linked to the full-length sequence of SEQ ID NO:15.

10

27. A method of treating individuals with prostate cancer comprising:

- a) identifying a regulatory protein that specifically binds to the promoter of prostate specific transglutaminase;
- b) identifying an activator or inhibitor of said regulatory protein; and
- 15 c) providing to an individual with prostate cancer an effective amount of said activator or inhibitor.

28. The method of claim 27, wherein said regulatory protein is identified by searching the sequence of SEQ ID NO:15 for sequences homologous with DNA-binding sites for known  
20 regulatory proteins.

29. The method of claim 28, wherein said regulatory protein binds to a site selected from a group consisting of GATA1, LMO2COM, CEBPB, XFD1, SRY, HFH2, HNF3B, RFX1, CMYB, ATF, NF1, S8, AP1, MZF1, AP2, BRN2, SP1, SRF, PADS, AP4, CAC binding, GFI1,  
25 OCT1, GR, TFIID, c-Myc, RORA1, VBP, AP4, IK2, Sp1, Tjian GC Box and CP2.

30. A method of treating individuals with prostate cancer comprising:

- a) providing a eukaryotic expression vector, said vector comprising a prostate specific transglutaminase promoter operably linked to a selected gene; and
- 30 b) providing to an individual with prostate cancer an effective amount of said vector.

31. The method of claim 30, wherein said gene encodes prostate specific transglutaminase, cytokeratin 15 or semenogelin II.
32. The method of claim 30, wherein said gene encodes thymidylate kinase, p53, cytosine  
5 deaminase, PNP, fibroblast growth factor receptor 2 (FGFR2), nitroreductase, PTEN, FHIT, KAI1 or diptheria toxin.
33. The method of claim 30, wherein said gene encodes a cytokine.
- 10 34. The method of claim 30, wherein said gene encodes an anti-bcl-2 ribozyme, an anti-sense c-myc RNA, an anti-sense cathepsin D RNA, E-cadherin, or polyglutamine.
35. The method of claim 30, wherein said gene encodes hepatitis B surface antigen, herpes  
simplex glycoprotein, HIV envelope proteins, hemagglutinin, malaria circumsporozoite protein,  
15 carcinoembryonic antigen, prostate-specific membrane antigen (PSMA), prostate stem cell antigen (PSCA), caveollin, POV1, HER2/neu or p27KIP1.
36. The method of claim 30, wherein said gene encodes a humanized antibody.
- 20 37. The method of claim 36, wherein said humanized antibody binds specifically to PSMA, PSCA, caveolin, POV1, HER2/neu or p27KIP1.
38. The vector of claim 5, wherein said gene encodes a tumor suppressor, a cytokine, a  
receptor, an oncogene or an inducer of apoptosis.
- 25 39. The vector of claim 38, wherein said gene is a proapoptotic Bcl2 gene selected from the group consisting of Bax, Bak, Bim, Bik, Bid, and Bad.
40. The vector of claim 38, wherein said tumor suppressor is selected from the group  
30 consisting of p53, p16, p21, MMAC1, p73, zac1, BRCA1 and Rb.
41. The vector of claim 38, wherein said inducer of apoptosis is selected from the group consisting of Harakiri, Ad E1B and an ICE-CED3 protease.



42. The vector of claim 38, wherein said cytokine is selected from the group consisting of IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, TNF, GMCSF,  $\beta$ -interferon and  $\gamma$ -interferon.

5

43. The vector of claim 38, wherein said receptor is selected from the group consisting of CFTR, EGFR, VEGFR, IL-2 receptor and the estrogen receptor.

44. The vector of claim 38, wherein said oncogene is selected from the group consisting  
10 *ras, myc, neu, raf, erb, src, fms, jun, trk, ret, gsp, hst, and abl.*

1/5

GCCTAGGCAT	TCCAGACCAG	CCTGGGCAAT	GCAGTGATAT	GCTCTCATTG	-1351
TGTAAATAAA	ACAAAATAAT	ACATTATTTG	CTTTGCAGAT	GCCACTGCCG	-1301
CCAGGAGCCC	TGTAACATCA	GCCATCGTGT	ATCCCACCGT	GTTCTTCAAC	-1251
GTCACCGTTG	GCGGCAAGCC	TTTAGACCTG	CGTCTCCTTC	AAGCTGTTTG	-1201
CAGACAAGGT	TCCAAAGCCA	GCAGAAAAC	TTTGTACTCT	GAGCACTGGA	-1151
GAGAAAGGAT	TTGGTTATAA	GAGTTCCTGC	TTTCACAGAA	TTATTCCAGG	-1101
GGTTATGTGT	CAGGGTGGTG	ACTTCATACG	CCATAATGGC	ACTGGTGGCA	-1051
AATCCATCTA	TGGGGAGAAA	TTTGATGATG	AGAACTTCAT	CCTAAAGCAT	-1001
ACAGGTCCTG	GCATGTTGTC	CATGGCAAAT	GCTGGACCCA	ATACAAATGG	-951
TTCCCAGTTT	TTAATCTGCA	CTGCCAAGAC	TGAGTGGTTG	GATGGCAAGC	-901
TTGTGGTCTT	TGGCAAGGTG	AAGGCATGAA	TATTGTGGAG	GCCATGGAGT	-851
GCTTTGTGTC	CAGGAATGGC	AAGACCGGCC	AGAAGATCAC	CATTGCTGAC	-801
TGTGGACAGC	TCTTATAAGT	TTGACTTGTC	TTTTATCTTA	ACCACCAGAC	-751
CATTCTTCT	GTAGCTCAGG	GGAGCACCT	CCACCCCAT	TGCTCCCAGT	-701
AACCTAGAAT	CTTTGTGCTC	TCGCTGCGGT	CCCCTCTGGG	GTGCCATGTT	-651
TTCTTGTTTC	CCTTCCATGC	CTAGCTGGAT	TGCAGAGTTA	AGTTTATGAT	-601
TATGAAAATA	AAAACCTAAA	TAACAAAATA	ATAATAATAG	TACATTATTT	-551
GCCTCTCAGT	GCATACAACA	TCTCTGGAAT	GATCTAAAAG	AAACTGGTAG	-501
TAGGATTCAA	CAAGGTGGAC	AGGGAGCATT	TTTTTCCATT	TTGAACCATG	-451
TGGATGCATT	ATATATTCAA	AATGTAACCA	AAAAAACTCA	CACGCATCTC	-401
TTACTTTTTTC	AACTTGAAGA	AAAGAGAAGA	AGAAGAAACC	CCTAAAGTCA	-351
GTAAAGGTTA	GCCTGGTTTC	CTAGACTTAA	GCTTGATAGT	AACCAAGAAT	-301
GTCAGGCCCA	CATGTGGTGT	AGGTCATAGG	GTCCTGGCTT	TTGGTCACCC	-251
ACTGTTAGGC	TCATATAACA	GACAAGGAGG	TGCCATTGTC	CGTTGTCTGT	-201
CCCTTTTACA	TCAGCTCATC	TTGGCTCCTG	ATTTCTCTGG	TCTGATATGG	-151
AAGATTCCGG	GATTTTATTA	TTTCTAATCA	ACATCTTTTT	CCCAAACCCC	-101
GCCCCTTGGC	AGCCATAGCA	AGGCCACAAA	AAAACCCAGG	ATCACATAAA	-51
AAGGGCTGTA	TAATCCCCTG	AGGACCGACT	GTGTGGAAGC	ACCAGGCATC	-1
+1 AGAGATAGAG	TCTACCCCTA	GCATTGCAGG	AGAGAATCTG	AAGAATCACT	+50
AGT					

FIG. 1

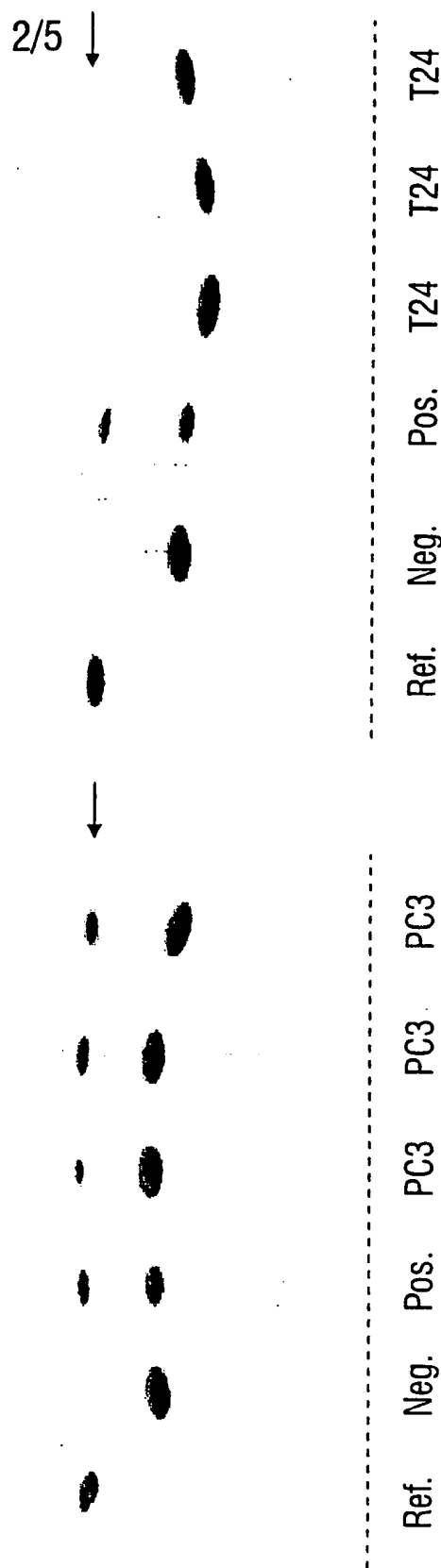


FIG. 2A

3/5

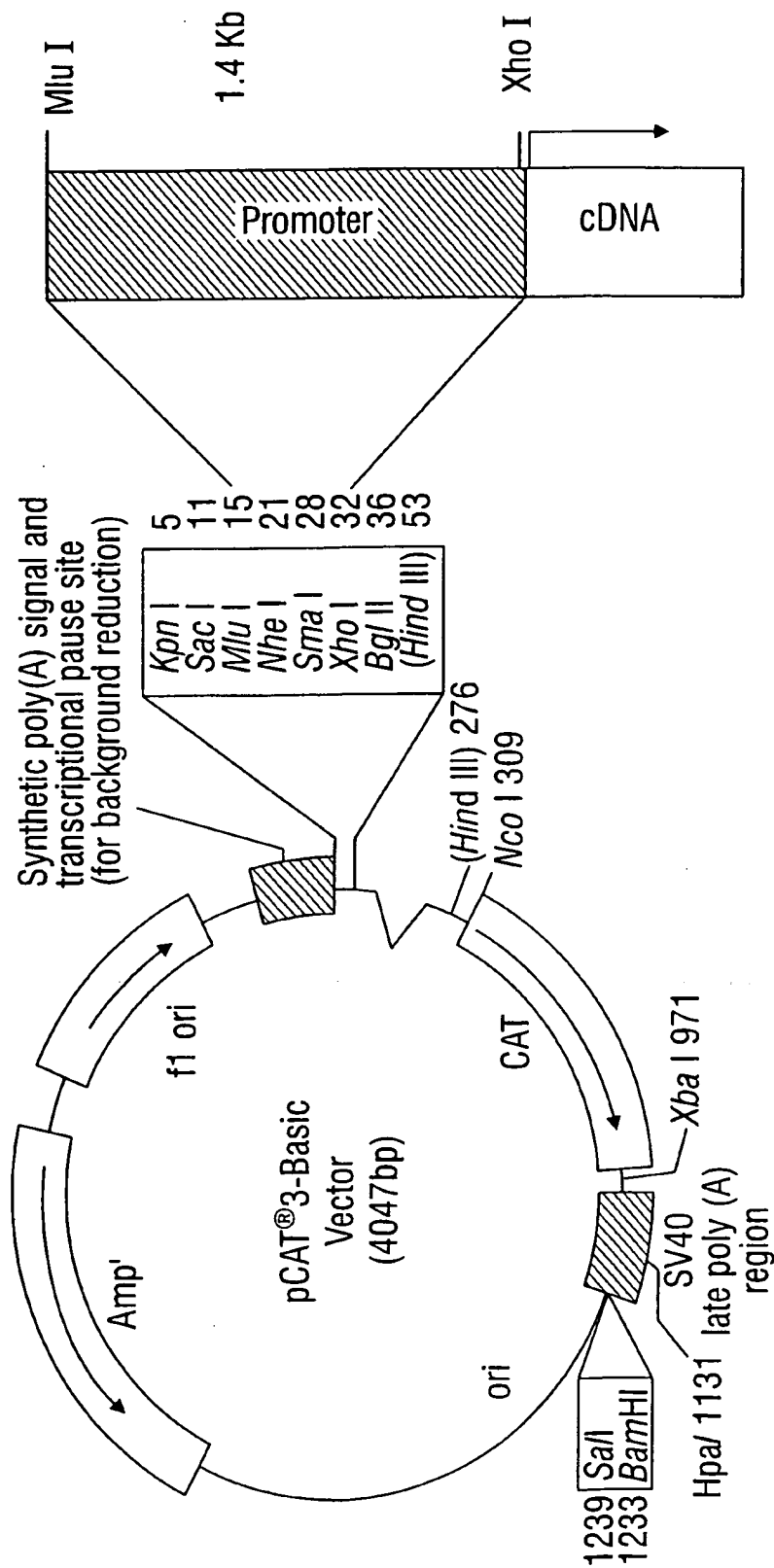


FIG. 2B



FIG. 3A

5/5

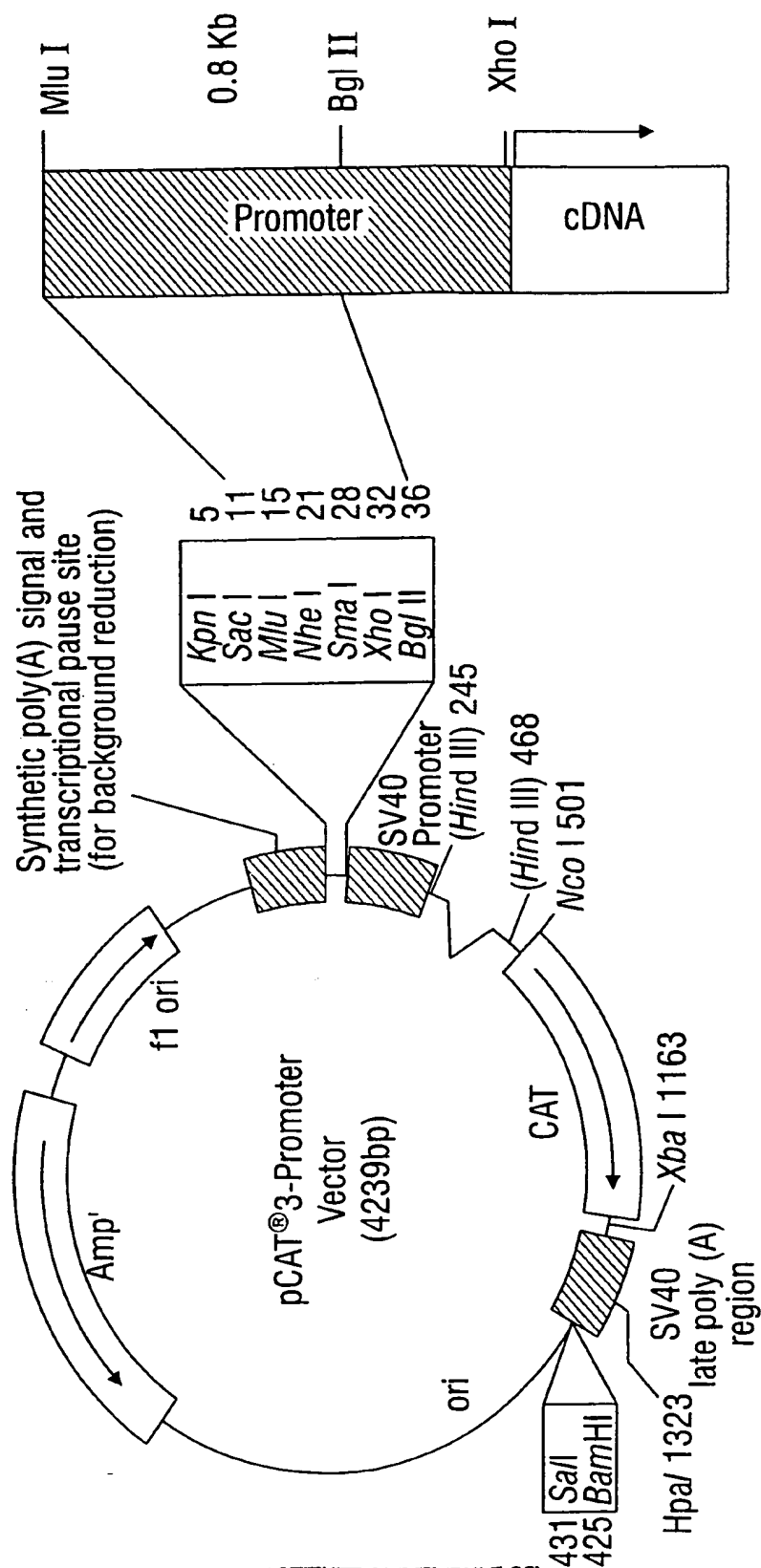


FIG. 3B

## SEQUENCE LISTING

<110> An, Gang  
Veltri, Robert W.

<120> PROSTATE SPECIFIC PROMOTER AND REGULATION OF GENE  
EXPRESSION

<130> UROC:020P

<140> Unknown

<141> 1999-09-07

<150> 60/099,338

<151> 1998-09-08

<160> 17

<170> PatentIn Ver. 2.0

<210> 1

<211> 3064

<212> DNA

<213> Homo sapiens

<400> 1

```
aattctaaaa atgcttttgc aagcttgcag gcctgcaggt gcagcggccg ccagtgtgat 60
ggatatctgc agaattcggc ttgcgctcag ctggaattcc gcagagatag agtcttccct 120
ggcatttcag gagagaatct gaagggatga tggatgcac aaaagagctg caagttctcc 180
acattgactt cttgaatcag gacaacgccg tttctcacca cacatgggag ttccaaacga 240
gcagtccctg gttccggcga ggacaggtgt ttcacctcg gctggtgctg aaccagcccc 300
tacaatccta ccaccaactg aaactggaat tcagcacagg gccgaatcct agcatcgcca 360
aacacaccct ggtggtgctc gacccgagga cgccctcaga ccactacaac tggcaggcaa 420
cccttcaaaa tgagtctggc aaagaggtca cagtggctgt caccagtcc cccaatgcca 480
tcctgggcaa gtaccaacta aacgtgaaaa ctggaaacca catccttaag tctgaagaaa 540
acatcctata ccttctcttc aacccatggt gtaaagagga catggttttc atgcctgatg 600
aggacgagcg caaagagtac atcctcaatg acacgggctg ccattacgtg ggggctgcca 660
gaagtatcaa atgcaaacc tggaaactttg gtcagtttga gaaaaatgtc ctggactgct 720
gcatttccct gctgactgag agctccctca agcccacaga taggagggac cccgtgctgg 780
tgtgcaggcg catgtgtgct atgatgagct ttgagaaaagg ccagggcgtg ctcatggga 840
attggactgg ggactacgaa ggtggcacag cccatacaa gtggacaggc agtgccccga 900
tcctgcagca gtactacaac acgaagcagg ctgtgtgctt tggccagtgc tgggtgtttg 960
ctgggatcct gactacagtg ctgagagcgt tgggcatccc agcacgcagt gtgacaggct 1020
tcgattcagc tcacgacaca gaaaggaacc tcacggtgga cacctatgtg aatgagaatg 1080
gcgagaaaaa caccagtatg acccacgact ctgtctggaa tttccatgtg tggacggatg 1140
cctggatgaa gcgaccctac gacggctggc aggcgttgga cgcaacgccg caggagcgaa 1200
gccaggggtg cttctgctgt gggccatcac cactgaccgc catccgcaaa ggtgacatct 1260
ttattgtcta tgacaccaga ttctgtctct cagaagtga tggtagacag ctcactctgt 1320
tggtagaatg ggtgaatggg caggaggagt tacacgtaat ttcaatggag accacaagca 1380
tcgggaaaaa catcagcacc aaggcagtgg gccaaagacag gcggagagat atcacctatg 1440
agtacaagta tccagaaggc tcctctgagg agaggcaggt catggatcat gccttctctc 1500
ttctcagttc tgagaggggag cacagacagc ctgtaaaaga gaactttctt cacatgtcgg 1560
tacaatcaga tgatgtgctg ctgggaaact ctgttaattt caccgtgatt cttaaaagga 1620
agaccgctgc cctacagaat gtcaacatct tgggctcctt tgaactacag ttgtacactg 1680
gcaagaagat ggcaaaactg tgtgacctca ataagacctc gcagatccaa ggtcaagtat 1740
cagaagtga tctgaccttg gactccaaga cctacatcaa cagcctggct atattagatg 1800
atgagccagt tatcagaggt ttcattcatt cggaatttgt ggagtctaag gaaatcatgg 1860
cctctgaagt attcacgtca aaccagtacc ctgagttctc tatagagttg cctaacacag 1920
gcagaattgg ccagctactt gtctgcaatt gtatcttcaa gaataacctg gccatccctt 1980
tgactgacgt caagttctct ttggaaagcc tgggcatctc ctactacag acctctgacc 2040
atgggacggg gcagcctggt gagaccatcc aatcccaaat aaaatgcacc ccaataaaaa 2100
```

```

ctggacccaa gaaatttatc gtcaagttaa gttccaaaca agtgaaagag attaatgctc 2160
agaagattgt tctcatcacc aagtagcctt gtctgatgct gtggagcctt agttgagatt 2220
tcagcatttc ctaccttgtg cttagctttc agattatgga tgattaaatt tgatgactta 2280
tatgagggca gattcaagag ccagcaggtc aaaaaggcca acacaacat aagcagccag 2340
acccacaagg ccaggctctg tgctatcaca gggtcacctc ttttacagtt agaaacacca 2400
gccgaggcca cagaatccca tccctttcct gagtcatggc ctcaaaaatc agggccacca 2460
ttgtctcaat tcaaattccat agatttcgaa gccacagagc tcttccctgg agcagcagac 2520
tatgggcagc ccagtgtctg cacctgtctg cgacccttga gaagctgcca tatcttcagg 2580
ccatgggttc accagccctg aaggcacctg tcaactggag tgctctctca gcactgggag 2640
gggcctgata gaagtgcatt ctctctctat tgctctcatt ctctctctc tatccctgaa 2700
atccaggaag tccctctcct ggtgtctcaa gcagtttgaa gccaatctg caaggacatt 2760
tctcaagggc catgtggttt tgcagacaac cctgtcctca ggctgaact caccatagag 2820
acccatgtca gcaaacggtg accagcaaat cctcttccct tattctaaag ctgccccttg 2880
ggagactcca gggagaaggc attgcttcct ccctgggtg aactctttct ttggtattcc 2940
atccactatc ctggcaactc aaggctgctt ctgttaactg aagcctgctc cttcttgctc 3000
tgccctccag agatttgctc aaatgatcaa taagctttaa attaaaccgg aatccgcgga 3060
attc 3064

```

&lt;210&gt; 2

&lt;211&gt; 1709

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

```

ggtacctcct gccagcacct cttgggtttg ctgagaactc acggggtcca gctacctggc 60
catgaccacc acatttctgc aaacttcttc ctccaccttt ggggggtggc caaccggagg 120
gggttccctc ctggctgggg gaggtggctt tgggtggggg agtctctctg ggggagggtg 180
aagccgaagt attctagctt cttctgctag gtttgtctct tcagggtcag gaggaggata 240
tggggggtggc atgagggctt gtggcttttg tggaggggct ggtagtgttt tcggtggagg 300
ctttggaggg ggcgttggtg ggggttttg tgggtggctt ggtggtggcg atggtggtct 360
cctctctggc aatgagaaaa ttaccatgca gaacctcaat gaccgccttg cctcctacct 420
ggacaaggta cgtgccttg aggaggccaa tgctgacctg gaggtgaaga tccatgactg 480
gtaccagaag cagacccag ccagcccaga atgcgactac agccaatact tcaagacct 540
tgaagagctc cgggacaaga tcatggccac caccatcgac aactcccggg tcatcctgga 600
gatcgacaat gccaggtctg ctgcggacga cttcaggctc aagtatgaga atgagctggc 660
cctgcgccag ggcgttgagg ctgacatcaa cggcttgccg cgagtccttg atgagctgac 720
cctggccagg actgacctg agatgcagat cgagggcctg aatgaggagc tagcctacct 780
gaagaagaac cacgaaggag agatgaagga gttcagcagc cagctggccg gccagggtcaa 840
tgtggagatg gacgcagcac cgggtgtgga cctgaccctg gtgctggcag agatgaggga 900
gcagtacgag gccatggcgg agaagaaccg ccgggatgtc gaggcctggt tcttcagcaa 960
gactgaggag ctgaacaaag aggtggcctc caacacagaa atgatccaga ccagcaagac 1020
ggagatcaca gacctgagac gcacgatgca ggagctggag atcgagctgc agtcccagct 1080
cagcatgaaa gctgggctgg agaactcact ggccgagaca gagtgccgct atgccacgca 1140
gctgcagcag atccaggggc tcatgtgttg cctggaggcc cagctgagtg agctccgatg 1200
cgagatggag gctcagaacc aggagtacaa gatgctgctt gacataaaga cacggctgga 1260
gcaggagatc gctacttacc gcagcctgct cgagggccag gatgccaaga tggtctggat 1320
tggtcatcagg gaagcctctt caggaggttg tggtagcagc agcaatttcc acatcaatgt 1380
agaagagtca gtggatggac aggtgggttc tccccacaag agagaaatct aagtgtctat 1440
tgcaggagaa acgtcccttg cactcccca ctctcatcag gccaagtgga ggactggcca 1500
gagggcctgc acatgcaaac tccagtcctt gccttcagag agctgaaaag ggtccctcgg 1560
tcttttattt cagggtcttg catgcgctct attcccctc tgctctccc caccttcttt 1620
ggagcaagga gatgcagctg tattgtgtaa caagctcatt tgtacagtgt ctgttcatgt 1680
aataaagaat tacttttctt ttgtgcaat 1709

```

&lt;210&gt; 3

&lt;211&gt; 1981

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

```

agacaagatt tttaagcaa gatgaagtcc atcatcctct ttgtccttcc cctgctcctt 60

```



```

atcttggaga agcaagcagc tgtgatggga caaaaagggtg gatcaaaagg ccaattgcca 120
agcggatctt cccaatttcc acatggacaa aagggccagc actatttttg acaaaaagac 180
caacaacata ctaaattccaa aggcagtttt tctattcaac acacatatca tgtagacatc 240
aatgatcatg actggaccgg aaaaagtcag caatatgatt tgaatgccct acataaggcg 300
acaaaatcaa aacaacacct aggtggaagt caacaactgc tcaattataa acaagaaggc 360
agagaccatg ataaatcaaa aggtcatttt cacatgatag ttatacatca taaaggaggc 420
caagctcatc atgggacaca aaatccttct caagatcagg ggaatagccc atctggaaa 480
ggattatcca gtcaatgttc aaacacagaa aaaaggctat gggttcatgg actaagtaa 540
gaacaagctt cagcctctgg tgcacaaaaa ggtagaacac aagggtggatc ccaaagcagt 600
tatgttctcc aaactgaaga actagtagtt aacaaacac aacgtgagac taaaaattct 660
catcaaaata aagggcatta ccaaaatgtg gttgacgtga gagaggaaca ttcaagtaa 720
ctacaaactt cactccatcc tgcacatcaa gacagactcc aacatggacc caaagacatt 780
tttactaccc aagatgagct cctagtatat aacaagaatc aacaccagac aaaaaatctc 840
agtcaagatc aagagcatgg ccggaaggca cataaaatat catacccgctc ttcacgtaca 900
gaagaaagac aacttcacca tggagaaaag agtgtacaga aagatgtatc caaaggcagc 960
atttctatcc aaactgaaga gaaaatacat ggcaagtctc aaaaccagggt aacaattcat 1020
agtcaagatc aagagcatgg ccataaggaa aataaaatat cataccaatc ttcaagtaca 1080
gaagaaagac atctcaactg tggagaaaag ggcatccaga aagggtgtatc caaaggcagt 1140
atttcgatcc aaactgaaga gcaaatacat ggcaagtctc aaaaccagggt aagaattcct 1200
agtcaagctc aagagtatgg ccataaggaa aataaaatat cataccaatc ttcgagtaca 1260
gaagaaagac gtctcaacag tggagaaaag gatgtacaga aagggtgtatc caaaggcagt 1320
atttctatcc aaactgaaga gaaaatacat ggcaagtctc aaaaccagggt aacaattcct 1380
agtcaagatc aagagcatgg ccataaggaa aataaaatgt cataccaatc ttcaagtaca 1440
gaagaaagac gactcaacta tggaggaaag agcacgcaga aagatgtatc ccaaagcagt 1500
atttctttcc aaattgaaaa gctagtagaa ggcaagtctc aaatccagac accaaatcct 1560
aatcaagatc aatgggtctgg ccaaaatgca aaaggaaagt ctggtcaatc tgcagatagc 1620
aaacaagacc tactcagtca tgaacaaaaa ggcagatata aacaggaatc cagtgaagtc 1680
cataatattg taattactga gcatgagggt gcccaagatg atcatttgac acaacaatat 1740
aatgaagaca gaaatccaat atctacatag ccctgttgct tagcaaccac ttgaaaagct 1800
ggaccaatag caagggtgtc cccgacctca gtgaagtctt tgatgtttct gagaggcaga 1860
ctcccatgtg gtcccatgac cttgggtccat ggatgacacc accttcccat gcttccttgc 1920
attaggcttt ctaaaccggg agccccttca aacttccaat aaagggatca ttttctgctt 1980
t

```

<210> 4  
 <211> 22  
 <212> DNA  
 <213> Homo sapiens

<400> 4  
 cgcctcaggc tggggcagca tt 22

<210> 5  
 <211> 25  
 <212> DNA  
 <213> Homo sapiens

<400> 5  
 acagtggaag agtctcattc gagat 25

<210> 6  
 <211> 25  
 <212> DNA  
 <213> Homo sapiens

<400> 6  
 cgagctgcct gacggccagg tcate 25

<210> 7  
 <211> 25  
 <212> DNA

<213> Homo sapiens

<400> 7

gaagcatttg cggtggacga tggag

25

<210> 8

<211> 24

<212> DNA

<213> Homo sapiens

<400> 8

gggggctgcc agaagtatca aatg

24

<210> 9

<211> 24

<212> DNA

<213> Homo sapiens

<400> 9

tgccaccttc gtagtcccca gtcc

24

<210> 10

<211> 23

<212> DNA

<213> Homo sapiens

<400> 10

tcttcaggag gtggtggtag cag

23

<210> 11

<211> 23

<212> DNA

<213> Homo sapiens

<400> 11

gagaggcaga gggggaatag agc

23

<210> 12

<211> 28

<212> DNA

<213> Homo sapiens

<400> 12

acatctcaac tgtggagaaa agggcatc

28

<210> 13

<211> 28

<212> DNA

<213> Homo sapiens

<400> 13

tgatcatctt gggcaacctc atgctcag

28

<210> 14

<211> 8224

<212> DNA

<213> Homo sapiens

<400> 14

gagctcaggc aatccacccg tcttggcctc ccaaagtgt aggattacag ccaccgtgcc 60  
cagcccgaat caatgcaatt cttatcaaaa attttatggc tttcttgtag aaatggaaaa 120

```

gtgggttcta acgttcattt ggaatttcaa gggatcccaa atagccgtac aatcttgact 180
aagaggaaca cagttggaag actcacactt cacaaatact atcttataac ccattattta 240
aactgacaac aacttaacac tgcttccata aacaaacaag agaaagaaaa ctaataaaga 300
ctctatactt taacttcatt cccgccactt tttaactgat aattgctgtg ctttctctca 360
ccccatgcac agaaatgctc tgtgccccat acctgcaacc gggagatgaa ggaggagtgg 420
catcagtgat acaagtgttt ttcttacctt tttagcgctt ctttcagtga tataaagtta 480
aaatcagtta ctgtgagtg ctcacctgatt ttttggttct tataaagggtg atctttctct 540
gcagatactt gttaaattgg tgaccttggg gggggcaggg gggatcattg gagccttcta 600
ttccaccatc ttgttctgcc acctctctct ctttaattttt attctctctg aacgatttta 660
tttctccttc atttctgaag gatagctttg ctgagggtaa tgttcttagc tgacagtttt 720
ctcttttagt aatttgaata tattatgcca ttctttcttg gcctgtaagg tttctcccga 780
gaaattctct gttagtcaga tgggggttac cttatatgtg acttgacatt tttctcttgc 840
tgcttttaaa aatctttctt tgtctttgac ttttagcaat gcaattaaaa tgtgccttgg 900
agagaacccg tttaggttga atttatttag ggttcttata gcttctctgga cctggatttc 960
tatctctctc ccaagatgca agaagtttt tgctattatt ttatgctct gatggaaaa 1020
ttgcctcttc caattttatg gagtagattt tataagcgaa aacttatttt tatgaatggg 1080
tctttagggt ttagtctatt tgagtttgtt ggttttgtt gcaggtgtat gtagcagtat 1140
tgtttccata gagtttcttc acctatgatc cacacttggt gcacttgaga gtttctcagc 1200
gccctagttg agagagtgtg ctgctacagt ggcattggct tgccagtgtc gtgggtccag 1260
cgctaattct caggtcaggg gtgtgtgcat ccacgtggta ggtaattcaa ttaagggtatg 1320
actcaatggt attggggcca ggggtctgtt atcttggtg gaagcatggg cacatgtttg 1380
ttcagtcagt gagtgtgact acaggctagg ggtggcctgt gaagctgttt ctgtgtttca 1440
ggcagtgga tctacagctt tcaactgggt tgaagacatg tctgccaggg tcagtctact 1500
gggcttttct tcaggcccaa gatgtgggtg taaggctgct tggatggcct acgagaggaa 1560
gggaatgaga gaggaagag ggctttttct caggcacttg acacagctgc atggctgagt 1620
ggctaggctt tgagcatttc tgtaaggggt ggccaagaa gatattctgt aggcctaggg 1680
tgtaggcagt gagcttcttg gttgtccttc atgtgtgacc accagaggca atctgttaga 1740
ctgtttctca gacctgggac atgaccacat aactacttgg ctggcttggg tatgcccacc 1800
agggttggtc caatgggtgt ttctcaggtt tgggaaatgg gctctcggca gctggggccag 1860
cgtgggtgtc agcccagcag gggagcttgt gagtctgttt ctcaagctct tattgggagt 1920
atagggccac tgagcaggcc atggatgtgt ctctggaggg agagagtgtt acagggctat 1980
ttctcaggtc ctacgtgttg acacatagcc actgtgctgc cacaggggca tataaactgt 2040
gtcaagctca gtgacctgtc tctacttgg tgaactttaa aatattctac tattttttcc 2100
ttgtttattt tttatcattt gatattttcc caatcatgat accacatcta agatgtatgc 2160
agcactatac catgtttatg ctttccaaat taccgatttc tgccttcac tttatcaata 2220
gatttatttt taattttgaa atgctttttt ttcaaatctc attaaatcca atcatgatat 2280
gagtatcata atcacatgtt gtgcctagtt tgcattttcc tgccttgttt ttgcaaactc 2340
tgtgcacctt tgttttatct gtatcttttg gacactagat tagttggaag ttttgttctt 2400
gatccataag taagattaat atttcagact ctgttttgtt ccggtaatct tggaaacaaa 2460
cagttctctg tatttctcta agggaaagtgt aaggaaagatt atagtgatta gcaggtacag 2520
caataaattt gtgttcagga aaactttact ctttctttt ggggtccctg ccataaggaa 2580
atacttactg cccttagggc aaacagataa aactggtaag atctcagaat ggctcataga 2640
gaaaccacat agctcacatg cttcaagttt gatggcataa gtagcatttt ttcttatggg 2700
tggtctttcc cctctcccct tccatatagc attgtttaca ggaagggttag gattttaaga 2760
gagaataata gaaaataagt ttccaagccg tcctttctcc agaactacct tccacagtgt 2820
ctgtaatgca tacagggtca ctgttcaa atgtgtgctt ctctgttgtt cttctcataa 2880
caaggacttt tttagtcaaa gtaaagaaac tggcgggtg cagtggctca tgcctgtaat 2940
cccagcagtt tgggaggtg aggcggggcg atcacagagt caggagattg agaccatcct 3000
ggctaacacg gtgaaacccc gtcttacta aaaatacaaa aaattagccg ggcgtgggtg 3060
caggcacctg tagtcccagc tacttgggaa gctgaggcag gagaatggca tgaacctggg 3120
aggcacagct tgacgtgagc cgagatgggt ccactgcact ccagcctggg cgacagagcg 3180
agactctgtc tcaaaaacaa acaaaacaaa aaaacaaaac aaacaaacaa acaaaacaaa 3240
aaccttctct ggatttttca agaagttttc acaattgcaa aaaaacaaaa caaaacaaa 3300
aacacaaca aacacaaaaa aaagcaaaaa aacctatgc ttccaccaa ggtaaaaatt 3360
ttagctttag gtccactctc aatacattat ttaatgaact gaagttggca aatattctct 3420
acagcctgtt ggagggttca gcagtttatt acagaagtat gaaatgcttt tatttaaaaa 3480
atgtattttg gtaaatatcat ttttgttttag gtaaatatca tcaataatc caatttggaa 3540
accaacattc ttacttcttt ttccaacagt tgttctatc atcataaaaa catgttaaat 3600
tttctctcat cttcaaaaa atctctggag cttacctcac cctccagata cagcctcacc 3660
tctcattcca ccatgaaatc agacttcttg agatgggttc agctggactc catccttcaa 3720
ctgacaaccc aacacaactg tattcatctc tegttaacat tactagcagt gagtaacatc 3780

```

```

agaaagtttt tgacacatgt tagtcttttt tgtgatgaac ttcacagata catttgacat 3840
tggtatgcct catttatttg ttgaaatttt ttcttttggc ttccatgaag tttctttctc 3900
ttctgtatca ttctacttct atgactgctc ttctctcgagt aaaacagaat gtgtctcagg 3960
attactttta aacaagacaa agtatagagt tatacctaaa atttagtatt taagtatttg 4020
gatcagaaag gaaactcgca tttagagtat gaaggcattg tcagccacca attacttttg 4080
taacctgaag ctagtctcct tctactccg gactgaattt cttctgtata atgcaagcga 4140
tctggcatga tgatatacaa agaccgataa aatttttgctg gggattctga aagtaaaaaa 4200
aattgccttt gatattatgt ccccatgcta agtccctggg gactttgaca ttatccccca 4260
ctgagcaggg gtgaggaagc tggcatttac taataagcta tgaaagggca gtgccttttg 4320
acatttcagc tccaccataa gcacaccac tcaaggaaca tataaatgaa gagatccgct 4380
cagttctcag acaagatttt tcaagcaaga tgaagtccat catcctcttt gtcttttccc 4440
tgctccttat cttggagaag caagcagctg tgatgggaca aaaagggtgag tggagagggt 4500
aagccttggg gaaagctact ttaaaaaaat ggcctctaag gatattcagg gtgcaaacag 4560
taacctgttc aggcacagat tcttctcctt gatgagaatt gatttttctc caccacaacgc 4620
tgtaggcttt tggaaatatt agaaatttgt tgggaaaagg tgggaggtaa gatttgcaag 4680
agagcttttg agataatgaa tgcatacatt tctattatca attaccagg gtatcaaaag 4740
gccaattgcc aagcgatct tcccaatttc cacatggaca aaagggccag cactattttg 4800
gacaaaaaga ccaacaacat actaaatcca aaggcagttt ttctattcaa cacacatatc 4860
atgtagacat caatgatcat gactggaccc gaaaaagtca gcaatatgat ttgaatgccc 4920
tacataaggc gacaaaatca aaacaacacc taggtggaag tcaacaactg ctcaattata 4980
aacaagaagg cagagaccat gataaatcaa aaggtcattt tcacatgata gttatacatc 5040
ataaggagg ccaagctcat catgggacac aaaaatcctt tcaagatcag gggaatagcc 5100
catctggaaa gggattatcc agtcaatgtt caaacacaga aaaaaggcta tgggttcattg 5160
gactaagtaa agaacaagct tcagcctctg gtgcacaaaa aggtagaaca caagggtgat 5220
cccaaagcag ttatgttctc caaactgaag aactagtagt taacaaacaa caacgtgaga 5280
ctaaaaatc tcatcaaaat aaagggcatt accaaaatgt ggttgacgtg agagaggaac 5340
attcaagtaa actacaaact tcactccatc ctgcacatca agacagactc caacatggac 5400
ccaaagacat ttttactacc caagatgagc tctagtata taacaagaat caacaccaga 5460
caaaaaatct cagtcaagat caagagcatg gccggaaggc acataaaata tcatacccg 5520
cttcacgtac agaagaaaga caacttcacc atggagaaaa gagtgtacag aaagatgtat 5580
ccaaaggcag catttctatc caaactgaag agaaaatata tggcaagtct caaaaccagg 5640
taacaattca tagtcaagat caagagcatg gccataagga aaataaaata tcataccaat 5700
cttcaagtac agaagaaaga catctcaact gtggagaaaa gggcatccag aaagggtgat 5760
ccaaaggcag tatttcgac caaactgaag agcaaatata tggcaagtct caaaaccagg 5820
taagaattcc tagtcaagct caagagtatg gccataagga aaataaaata tcataccaat 5880
cttcgagtac agaagaaaga cgtctcaaca gtggagaaaa ggatgtacag aaagggtgat 5940
ccaaaggcag tatttctatc caaactgaag agaaaatata tggcaagtct caaaaccagg 6000
taacaattcc tagtcaagat caagagcatg gccataagga aaataaaatg tcataccaat 6060
cttcaagtac agaagaaaga cgactcaact atggaggaaa gagcacgcag aaagatgtat 6120
cccaagcag tatttcttct caaattgaaa agctagtata aggcaagtct caaatccaga 6180
caccaaattc taatcaagat caatggtctg gccaaaatgc aaaaggaaag tctggtcaat 6240
ctgcagatag caaacaagac ctactcagtc atgaacaaaa aggcagatac aaacagggaat 6300
ccagtgagtc acataatatt gtaattactg agcatgaggt tgccaagat gatcatttga 6360
cacaacaata taatgaagac agaaatccaa tatctacata gccctgttgc ttagcaacca 6420
cttgaanaagc tggaccaata gcaaggtaag ttgtcttttc ttaccaaata ggagagggtgc 6480
ctgtcccaaa gttggggact ctccaggaac atggtaggac tgataaccat tgttcacatc 6540
aatagaagtg ctatattaca agtggtgagg agatgaacac catttctctg cgagtagagg 6600
acctggtagt ggcaggggaag gctgcttggc ctatcactgg gtcttagaat tcctattctt 6660
aattgagtat tcttcaataa tatttttata catgcctacc tgctaaagat ttttttgaac 6720
atgcactgac tatatatgca tatttatgag tttatgggat actcttgtca attcttatac 6780
tttagattag taaacctcaa attctttctc atatagtatg aaatattaca gcagttataa 6840
tttcttttct gcacatacat gaatgttctt gcatccctgt tagagttcat ctatgtctct 6900
tcagagacca caagcccaaa gactagcagt ccactctctc tgaatatagg aaagatatga 6960
gtagaaagaa ggattcctgt tcagatttggt aaaaaggaag tggaaatgga ggtgcaggag 7020
atgctgagag atctcagtg ctaactggac acttgcaatg tcaggagagg aaaattttta 7080
cctggatttg ggaaatggtt cttccatgcc cctttgcaat aagtaacact gtacctgaag 7140
agagggtgag cagcagaac aatggtccca gctgataact agtgacctgg tgcctaatg 7200
atcagggggc tgggtgcagtt gacgctaaag gggacagggg tcccagctct cccatctca 7260
ccccactct ccactatct cactatctg ttgtctttt ttctccctag gtgtcaccgc 7320
acctcagtga agtctttgat gtttctgaga ggcagactcc catgtggtcc cagatccttg 7380
gtccatggat gacaccacct tcccatgctt ccttgcatga ggctttctaa acccgagacc 7440

```

```

ccttcaaact tccaataaag ggatcatttt ctgcttttate tgettttggc tccagtgate 7500
tctgaattcc tagtggtcctc agggggcgtga ggtcatttctc gataaaaaatg gttttggag 7560
aaaagaatgg gatattttttt aagcattagg gagaatagca atcaaagcca cattttctaga 7620
tgtattgata gaaatccaac actaaaaaca agatgctatt actctgccag gcacacatct 7680
acttcctgat gccaaccccta aagtctctcg tgacctctgt actaacagga ccccaggatc 7740
taacagtcctc ccagaaagaa cagagaaaat gtattgcaac aagccctagg aaagtccctg 7800
atctactagt gccatgagta ttccccaagt ccacagcagc atatctcttc acttagctta 7860
agaactgtcc tactectaca gcatacattt agatgcaatc acaccaaaga gtgatattgc 7920
aagtctccac cacttcagca agagtgagtg gatttaaaaa tgttatctgt gatggaaa 7980
ccagacgacc aggccagagt cttggggcag gtgggtgtgt aataatagaa attctactcc 8040
ttttatgccc ttccaccctt tattctaagg atgattccag ggaaagccaa acatccagga 8100
gaagaagcca acatagagga ataagaaaaa tcaactagctc aggaataaga gatctggatt 8160
ccattctagc tctgttttaa gaactagtat atctttaatt aaatcttttc tttcctgcaa 8220
gctt 8224

```

&lt;210&gt; 15

&lt;211&gt; 1453

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 15

```

gcctaggcat tccagaccag cctgggcaat gcagtgate gctctcattg tgtaaataaa 60
acaaaataat acattatttg ctttgcagat gccactgccg ccaggagccc tgtaacatca 120
gccatcgtgt atcccaccgt gttcttcaac gtcaccgttg gcggcaagcc tttagacctg 180
cgtctccttc aagctgtttg cagacaaggt tccaaagcca gcagaaaact tttgtactct 240
gagcactgga gagaaaggat ttgggtataa gagttcctgc ttccacagaa ttattccagg 300
ggttatgtgt caggggtggg acttcatacg ccataatggc actgggtggc aatccatcta 360
tggggagaaa tttgatgatg agaacttcat cctaaagcat acaggctctg gcatgtgtgc 420
catggcaaat gctggaccca atacaaatgg ttcccagttt ttaatctgca ctgccaaagac 480
tgagtgggtg gatggcaagc ttgtggtctt tggcaagggtg aaggcatgaa tattgtggag 540
gccatggagt gctttgtgtc caggaatggc aagaccggcc agaagatcac cattgctgac 600
tgtggacagc tcttataagt ttgacttgtg ttttatctta accaccagac cattccttct 660
ttagctcagg ggagcaccct ccacccatt tgetcccagt aacctagaat ctttgtgctc 720
tcgctgcggt cccctctggg gtgccatgtt ttcctgttc ccttccatgc ctagctggat 780
tgcagagtta agtttatgat tatgaaaata aaaacctaaa taacaaaata ataataatag 840
tacattattt gcctctcagt gcatacaaca tctctggaaa gatctaaaag aaactggtag 900
taggattcaa caagggtggc agggagcatt tttttccatt ttgaaccatg tggatgcatt 960
atatattcaa aatgtaacca aaaaaactca cagcatctc ttactttttc aacttgaaga 1020
aaagagaaga agaagaaacc cctaaagtca gtaaaagtta gcctggtttc ctgacttaa 1080
gcttgatagt aaccaagaat gtcaggccca catgtgtgt aggtcatagg gtcctggctt 1140
ttggtcaccc actgttaggc tcatataaca gacaaggagg tgccattgtc cgttgtctgt 1200
cccttttaca tcagctcatc ttggctcctg atttccctggg tctgatattg aagattccgg 1260
gattttatta tttctaatac acatcttttt cccaaacccc gcccttggc agccatagca 1320
aggccacaaa aaaaccagg atcacataaa aagggtctg taatcccctg aggaccgact 1380
gtgtggaagc accaggcatc agagatagag tctacccta gcattgcagg agagaatctg 1440
aagaatcact agt 1453

```

&lt;210&gt; 16

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 16

gggggctgcc agaagtatca aatg

24

&lt;210&gt; 17

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 17

WO 00/14234

PCT/US99/20544

8

tgccaccttc gtagtcccca gtcc

24

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/20544

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C12N15/55 C12N9/10 C12Q1/68 C07K14/47  
A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12Q C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DUBBINK, HENDRIKUS J. (1) ET AL: "Human prostate - specific transglutamine gene (TGM4) an promoter characterization." EUROPEAN UROLOGY, (SEPT., 1998) VOL. 34, NO. 3, PP. 255-256. MEETING INFO.: 13TH CONGRESS OF THE EUROPEAN SOCIETY FOR UROLOGICAL ONCOLOGY AND ENDOCRINOLOGY INNSBRUCK, AUSTRIA OCTOBER 1-3, 1998, XP000874269 abstract  ----- -/-	1-7, 15-17, 19,23, 24,26

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

7 February 2000

Date of mailing of the international search report

23/02/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Kania, T

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/20544

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DUBBINK, HENDRIKUS J. (1) ET AL: "The human prostate-specific transglutaminase gene (TGM4): Genomic organization, tissue-specific expression, and promoter characterization." GENOMICS, (AUG. 1, 1998) VOL. 51, NO. 3, PP. 434-444. , XP000874237 the whole document	1-44
A	WO 95 19434 A (CALYDON INC) 20 July 1995 (1995-07-20) the whole document	5-24, 26, 30-44
A	WO 96 26272 A (HESTON WARREN D W ;SLOAN KETTERING INST CANCER (US); ISRAELI RON S) 29 August 1996 (1996-08-29) see the whole document, esp. examples 2,5	5-24, 26, 30-44
A	WO 96 14875 A (UNIV CALIFORNIA) 23 May 1996 (1996-05-23) the whole document	5-24, 30-44
A	WO 94 03594 A (UNIV MANITOBA ;MATUSIK ROBERT J (CA)) 17 February 1994 (1994-02-17) see the whole document, esp. pp.18-24	5-24, 30-44
A	WO 97 35993 A (UNIV CALIFORNIA) 2 October 1997 (1997-10-02) the whole document	5-24, 30-44
A	SEGAWA T ET AL: "Prostate-specific amplification of expanded polyglutamine expression: a novel approach for cancer gene therapy." CANCER RESEARCH, (1998 JUN 1) 58 (11) 2282-7. , XP002129868 cited in the application the whole document	5-22, 30-44
A	LEE C -H ET AL: "PROSTATE-SPECIFIC ANTIGEN PROMOTER DRIVEN GENE THERAPY TARGETING DNA POLYMERASE-ALPHA AND TOPOISOMERASE II ALPHA IN PROSTATE CANCER" ANTICANCER RESEARCH, GR, ATHENS, vol. 16, 1 August 1996 (1996-08-01), pages 1805-1811, XP002912271 the whole document	5-22, 30-44
A	EP 0 717 105 A (PASTEUR INSTITUT) 19 June 1996 (1996-06-19) see pref. pp.11-13	25, 26
	-/-	



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/20544

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BOULIKAS T: "Gene therapy of prostate cancer: p53, suicidal genes, and other targets." ANTICANCER RESEARCH, (1997 MAY-JUN) 17 (3A) 1471-505. REF: 236 , XP000874263 the whole document	1-44
A	DUBBINK H. ET AL.: "Tissue-specific and androgen-regulated expression of human prostate-specific transglutaminase" BIOCHEMICAL JOURNAL, vol. 315, 1996, pages 901-908, XP002129869 the whole document	1-44
T	AN G ET AL: "Human prostate - specific transglutaminase gene: Promoter cloning, tissue-specific expression, and down-regulation in metastatic prostate cancer" UROLOGY, (DEC 1999) VOL. 54, NO. 6, PP. 1105-1111. PUBLISHER: ELSEVIER SCIENCE INC, 655 AVENUE OF THE AMERICAS, NEW YORK, NY 10010. ISSN: 0090-4295., XP000874251 the whole document	1-44

# INTERNATIONAL SEARCH REPORT

international application No.

PCT/US 99/20544

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claims 27-37  
are directed to a method of treatment of the human/animal  
body, the search has been carried out and based on the alleged  
effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such  
an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all  
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment  
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report  
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is  
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/20544

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9519434 A	20-07-1995	US 5830686 A	03-11-1998
		AU 692837 B	18-06-1998
		AU 1686995 A	01-08-1995
		CA 2181073 A	20-07-1995
		EP 0755443 A	29-01-1997
		JP 9509049 T	16-09-1997
		US 5648478 A	15-07-1997
WO 9626272 A	29-08-1996	US 5935818 A	10-08-1999
		AU 5172596 A	11-09-1996
		CA 2212846 A	29-08-1996
		EP 0812356 A	17-12-1997
WO 9614875 A	23-05-1996	US 5919652 A	06-07-1999
		AU 4148996 A	06-06-1996
		EP 0793508 A	10-09-1999
WO 9403594 A	17-02-1994	AU 4694793 A	03-03-1994
		CA 2142181 A	17-02-1994
		EP 0656060 A	07-06-1995
		JP 7509608 T	26-10-1995
		US 5783681 A	21-07-1998
		US 5952488 A	14-09-1999
WO 9735993 A	02-10-1997	CA 2250183 A	02-10-1997
		EP 0889968 A	13-01-1999
EP 0717105 A	19-06-1996	AU 4037395 A	20-06-1996
		CA 2165098 A	15-06-1996
		JP 8242866 A	24-09-1996
		NZ 280674 A	22-09-1997